

## **Кафедра комп'ютерної інженерії та електроніки**

Назва дисципліни – Статистична радіофізика

Викладач: Котик Михайло Васильович

1. Вступ до статистиної радіофізики
2. Випадкові процеси.
3. Випадкові поля
4. Теорія інформації і кодування.
5. Основи теорії інформації та ідеальне кодування.
6. Основи теорії кодування

связи, хотя и имеет некоторые интересные точки соприкосновения со статистической физикой (соотношение между понятиями информации и энтропии), носит в основном математический характер. Это, если можно так выразиться, «радиоматематика», а не радиофизика, как она очерчена выше.

Что касается второго из названных направлений радиофизики — применения радиометодов в самых различных областях физики, астрономии и т. д., то здесь, в той мере, в какой используются элементы радиосвязи (генерация, распространение, прием), могут сохранять свое значение те статистические явления, с которыми сталкиваются в области радиотехники, но, сверх того, могут добавиться и другие, характерные для данной конкретной области (оптические флуктуации, акустическая реверберация и т. д.).

Как правило, случайные явления, с которыми приходится иметь дело в радиофизике, — это *процессы*, протекающие во времени, или еще шире — *поля*, зависящие и от времени, и от точки пространства. Адекватным математическим аппаратом для трактовки таких явлений служит *теория случайных функций*, которая за последние десятилетия приобрела большое значение во многих областях физики и техники. Эта теория представляет собой непосредственное обобщение классической теории вероятностей, рассматривающей случайные события и величины. Если не углубляться пока в вопрос о том, что означает *задание* случайной функции, то можно сказать, что это такая функция  $f(t)$ , значение которой при каждом из возможных значений аргумента  $t$  есть случайная величина. Для простоты здесь взят единственный аргумент  $t$ , под которым в приложениях чаще всего приходится понимать время. Конечно, это еще не предreshает того, дискретны или непрерывны возможные значения  $t$ , чему соответствуют два типа случайных функций — случайные *последовательности* (дискретное  $t$ ) и случайные *процессы* (непрерывное  $t$ ).

Если аргумент  $t$  считать временем, то соотношение между классической теорией вероятностей и теорией случайных функций представляется аналогичным (в математическом аспекте) соотношению между статикой и динамикой в механике или между термостатикой и кинетикой в термодинамике. Основы классической теории вероятностей предполагаются в данном курсе уже известными.

Все сказанное позволяет пояснить теперь те соображения, исходя из которых было выбрано содержание и построение данного курса лекций.

С одной стороны, как это видно уже из того сжатого перечня проблем, который был приведен выше, область статистической радиофизики чрезвычайно обширна. Отразить в курсе все мно-

гообразие статистических задач радиофизики просто невозможно, да и вряд ли подобная попытка могла бы быть как-либо обоснована. Вместе с тем широкий обзор по необходимости был бы беглым и уже поэтому принес бы мало пользы. С другой стороны, математический аппарат статистической радиофизики значительно шире и менее знаком, чем классическая теория вероятностей. Полное освещение теории случайных функций — это предмет самостоятельного математического курса.

При этих условиях представлялось наиболее целесообразным выделить немногие радиофизические вопросы, которые, во-первых, важны и интересны сами по себе и, во-вторых, позволят ознакомиться с теорией случайных функций и ее применениями. Разумеется, речь идет при этом не о математическом изложении теории.

Хорошо известно, насколько трудно найти компромисс между требованиями, которым должно удовлетворять строгое изложение математической теории, и необходимой при первом соприкосновении наглядностью. Безукоризненная аргументация математика зачастую представляется физику или инженеру не более, а менее убедительной, чем простые и осязаемые соображения, которые в свою очередь, с точки зрения математика, лишены какой бы то ни было доказательной силы. Необходимо подчеркнуть поэтому, что данный курс написан не математиком и не для математиков. В своей математической части он дает лишь некоторую предварительную подготовку, достаточную для того, чтобы приступить к самостоятельной работе, но, конечно, не заменяющую углубленного изучения математических работ.

До главы III теория случайных функций вообще не затрагивается и речь идет о некоторых, настолько простых задачах, что при их трактовке можно обойтись средствами классической теории вероятностей.

## Г л а в а I

---

### ЗАДАЧА БЕРНУЛЛИ

#### § 1. Физическое понятие вероятности

Как показывает опыт преподавания, изучение теории вероятностей в чисто математическом плане не всегда приводит к достаточной ясности в принципиальных вопросах, существенных для физика. Целесообразно поэтому хотя бы коротко остановиться на том, что такое вероятность и как пользуется этим понятием физик.

Хорошо известно, что вычислять вероятности научились задолго до того, как по-настоящему поняли, что такое вероятность. Теория возникла из попыток рассчитать шансы на выигрыш, «справедливые» ставки и т. п. в так называемых азартных играх<sup>1)</sup>, примерами которых могут служить игра в кости, в орлянку, такие карточные игры, как «очко», т. е. игры, в которых все зависит от «воли случая» и ничего не зависит от способностей игрока, его сообразительности или умения. В этом смысле подобные игры можно назвать глупыми. Между тем, как любят говорить математики, умная игра в шахматы практически ничего не дала науке, тогда как глупая игра в кости дала очень много. В чем причина этого интересного факта?

Дело в том, что шахматы подчинены своим специальным правилам и все ситуации, которые здесь возникают, не выходят за рамки этих правил, т. е. не могут претендовать на какую-либо всеобщность. Напротив, игра в кости элементарно проста и позволяет проявиться в чистом виде чрезвычайно общей *статистической* закономерности — устойчивости относительных частот при возрастании числа испытаний. Если при  $n$  бросаниях кости число очков  $i$  выпало  $n_i$  раз, то относительная частота  $n_i/n$  обнаруживает с увеличением  $n$  удивительное постоянство. Этот

---

<sup>1)</sup> Французское слово «hasard» означает случай, риск.



эмпирический факт не зависит от того, «хорошая» кость или «плохая». У заведомо фальшивой кости, содержащей, например, кусочек свинца со стороны одной из граней, устойчивость величин  $n_i/n$  все равно имеет место, хотя для разных граней относительные частоты оказываются неодинаковыми, не равными  $1/6$ .

Итак, устойчивость  $n_i/n$  не требует, чтобы кость была «хорошей», а конкретные «асимптотические» значения  $n_i/n$  не вытекают из самого факта устойчивости. Их дает только *статистический опыт*, либо специально поставленный, либо накопленный ранее. Говоря, что у «хорошей» (пригодной для игры) кости  $n_i/n$  близки к  $1/6$ , мы, в сущности, определяем, какую кость мы будем называть «хорошей».

Иногда полагают, что значения  $n_i/n \approx 1/6$  следуют из принадлежащего Лапласу классического определения вероятности:

$$P = \frac{\text{число благоприятных исходов}}{\text{число равновозможных исходов}}.$$

Это определение недостаточно широко, так как оно не охватывает случаев, когда возможные исходы составляют бесконечное счетное или непрерывное множество. Но даже не требуя от определения больше того, на что оно распространяется, легко заметить тавтологичность данного определения. Ведь равновозможность означает здесь не что иное, как равновероятность, и, следовательно, вероятность  $P$  «определена» через вероятность же. Просто здесь дается правило подсчета вероятностей интересующих нас исходов по принятому заранее *равномерному* распределению вероятностей всех возможных исходов. Полагая, что из определения Лапласа следует вероятность  $P = 1/6$  выпадения каждой из граней кости, мы на самом деле лишь извлекаем из него то, что в него вкладываем, — равновозможность выпадения каждой из граней. Совершенно так же мы предполагаем, что равновозможно появление любой из 36 или 52 карт при извлечении одной карты из колоды и т. п.

На что опираются все такие «самоочевидные» предположения?

В качестве их опоры иногда привлекается так называемый «принцип недостаточного основания»: если кость сделана геометрически аккуратно, из однородного материала и т. п., то *нет оснований* считать иначе.

Психологически это, может быть, и понятно. Для человека, ровно ничего не знающего ни о каком статистическом опыте, равновероятность выпадения граней аккуратной кости или обеих сторон монеты представляется очевидной по «здравому смыслу». Конечное число возможных исходов и наличие *симметрии* делают здесь равновероятность естественной до самоочевидности. Но в отсутствие симметрии или, скажем, при непрерывных

возможных исходах «здравый смысл» пасует или подводит. В случае фальшивой кости только статистический опыт с *данной* костью позволяет получить оценку вероятностей выпадения различных граней. Трудности, возникающие при выборе равновероятных исходов при непрерывном их множестве, были очень ярко продемонстрированы на ряде геометрических задач в «Исчислении вероятностей» Бертрана (1888 г.). Эти так называемые парадоксы Бертрана с «геометрическими» вероятностями — еще один пример недостаточности «здорового смысла» или интуиции в этой более сложной ситуации [21]. Более того, когда мы говорим, что в простых случаях интуитивное предсказание *правильно*, то сама эта «правильность» означает лишь то, что проверка на статистическом опыте подтвердила бы предсказание. Таким образом, даже «самоочевидные» предположения о равновероятности в *конечном счете* опираются на огромное количество испытаний, фактически проведенных с соблюдением определенных условий, т. е. на накопленный статистический опыт.

Конечно, здесь нет возможности углубляться в историю развития теории вероятностей. На протяжении этой долгой истории возникали и сталкивались разные воззрения на понятие вероятности. Созревание теории вероятностей как аксиоматизированной ветви математики затянулось почти на три столетия.

Математическая теория, возникнув на основе каких-то (обычно довольно простых) идей, почерпнутых из реальных явлений и фактов, большей частью стремится в дальнейшем к эмансипации, к отрыву от своих эмпирических корней, к достижению уровня аксиоматизированной теории. У теории вероятностей этот процесс завершился лишь в 30-х годах нашего века, когда А. Н. Колмогоров сформулировал аксиомы, сделавшие теорию вероятностей главой метрической теории функций<sup>1)</sup>.

Сегодня математик называет вероятностью неотрицательную, нормированную к единице, вполне аддитивную функцию множеств, определенную на некоторой алгебре множеств. Событие  $A$  изображается множеством  $A$  точек пространства всевозможных «элементарных исходов» рассматриваемого опыта (испытания), и вероятность  $P\{A\}$  события  $A$  — функция множества такая, что если она имеет смысл для множеств  $A_1, \dots, A_n$ , то она может быть определена и для множества, состоящего из всех точек, входящих хотя бы в одно из множеств  $A_1, \dots, A_n$ , а также и для множества точек, входящих сразу во все множества

---

<sup>1)</sup> Следует отметить, что С. Н. Бернштейн уже в 1917 г. указывал на необходимость аксиоматизации теории вероятностей и предложил конкретный набор аксиом, оказавшийся, однако, заметно менее удобным, чем аксиомы, предложенные А. Н. Колмогоровым.

$A_1, \dots, A_n$  (см. [8, 22]). Эта функция подчинена следующим трем аксиомам:

I. Вероятность  $P\{A\}$  события  $A$  удовлетворяет неравенству  $P\{A\} \geq 0$ .

II. Для достоверного события  $U$  имеет место равенство  $P\{U\} = 1$ .

III. Для взаимно исключающих друг друга событий  $A_k$  ( $k = 1, 2, \dots, n$ , где  $n$  может быть сколь угодно велико)

$$P\left\{\sum_{k=1}^n A_k\right\} = \sum_{k=1}^n P\{A_k\}.$$

Знак суммы имеет различный смысл слева и справа. Слева он означает, что речь идет о вероятности того, что произойдет хоть какое-то из событий  $A_k$ , так что  $\sum_{k=1}^n A_k$  — это событие ( $A_1$  или  $A_2, \dots$ , или  $A_n$ ), а справа стоит обычная сумма неотрицательных чисел  $P\{A_k\}$ .

«Достоверное событие», о котором идет речь в аксиоме II, отвечает множеству  $U$ , состоящему из всех мыслимых «элементарных исходов», так что все другие события  $A$  представляют собой подмножества  $U$ .

Аксиома III (аксиома сложения) охватывает и случай  $n = \infty$ , в силу чего  $P$  и называется *вполне* аддитивной функцией.

Из аксиом I—III, дополненных определениями ряда связанных с вероятностью понятий (например, понятия случайной величины или понятия математического ожидания), логически вытекает вся теория вероятностей.

Конечно, в аксиомах, определяющих величину  $P$ , нетрудно разглядеть связь этой абстрактной величины с эмпирической относительной частотой, но это *генетическая* связь, касающаяся происхождения аксиом, а не их содержания, из которого все эмпирическое уже исключено. Поэтому, выслушав это абстрактное определение вероятности, физик, инженер, экономист и т. д., т. е. человек, имеющий дело с реальными вещами и явлениями, сразу же спросит, что ему делать с этой вполне аддитивной функцией множеств, как связывать уравнения и формулы математической теории с реальным миром. Ситуация здесь та же, что и во всякой физической теории.

Уравнения и формулы для неких величин, взятые сами по себе, еще не исчерпывают физической теории. Последняя требует, чтобы мы знали, как извлекать из реальных вещей и явлений те числа, которые следует подставлять в математические формулы в качестве значений входящих в них величин, т. е.

знали, как *измерять* эти величины<sup>1)</sup>. Разумеется, физическая теория представляет собой органическое целое, ее математическая и измерительная части никоим образом не независимы друг от друга, но они не заменяют одна другую. Располагая одной, нельзя обойтись без другой. Поэтому для всех физических (и вообще практических) приложений математической теории вероятностей необходимо дополнить последнюю по крайней мере одним (реализуемым и конкретным) *способом измерения* входящей в нее величины  $P$  — вероятности. Естественно обратиться в поисках такого способа, или «аксиомы измерения», к относительной частоте.

Примем, что вероятность события *измеряется* (приближенно, как и при любом измерении) *относительной частотой* его появления в достаточно длинной серии испытаний, осуществляемых при определенных неизменных условиях, в достаточно обширном ансамбле «однородных» систем, т. е. в статистическом опыте. Будучи самостоятельным постулатом, не содержащимся в аксиомах математической теории вероятностей, этот способ не предуказан как единственный и не дает априорных гарантий успеха. Конечно, от него надо заранее требовать, чтобы он был логически совместим с математической теорией. Как мы убедимся далее (§ 19), частотная «аксиома измерения» вероятности этому требованию удовлетворяет. Но приведет ли получающаяся в результате добавления этой аксиомы физическая теория случайных явлений к согласию с опытом — это уже *дальнейший* вопрос, на который может ответить только статистический опыт. Важно, однако, подчеркнуть, что без *какого-либо* способа измерения вероятностей нельзя не только ответить на подобный вопрос, но даже его поставить.

Принимая предложенный способ измерения вероятности, мы отнюдь не отождествляем ее с относительной частотой, как это делает *частотная концепция* вероятности, выдвинутая Р. Мизесом в 1928 г. [24]. Он предложил понимать под вероятностью предел относительной частоты  $n_i/n$  при  $n \rightarrow \infty$  в «статистическом коллективе», т. е. в ансамбле или в серии испытаний, удовлетворяющих некоторым требованиям. Однако такое определение понятия вероятности не могло дать удовлетворительное обоснование для математической теории. Ведь никто не знает, что такое предел эмпирической величины. Если же понимать «предел» в каком-либо вероятностном смысле (§ 19), то мы вновь оказываемся в порочном кругу, так как пытаемся определить понятие вероятности через вероятность.

<sup>1)</sup> Так, по крайней мере, ставит вопрос об измерениях классическая (доквантовая) физика. В уравнениях квантовой механики фигурируют и такие величины, непосредственное измерение которых не является необходимым. (См. в [23] Лекции по основам квантовой механики.)

Еще и поныне в зарубежной литературе пользуется иной раз успехом так называемая *субъективная концепция* вероятности, согласно которой вероятность есть мера нашего незнания. Мы не знаем, какая выпадет грань кости, и *поэтому*  $P = 1/6$ . Если принимать эту точку зрения всерьез, то происходит чудо: из самого незнания якобы рождается некое знание, некое положительное утверждение. В сущности, здесь утверждается, что если распределение неизвестно, то оно равномерно. Конечно, мы вправе испытать и такую гипотезу, с тем чтобы судить о ее пригодности по тем следствиям, к которым она приводит. Но почему именно эту гипотезу надо считать с необходимостью вытекающей из нашего незнания? В случае фальшивой кости мы тоже не знаем, какая выпадет грань, но наше незнание несколько не подвигает нас в установлении правильных значений вероятностей для граней такой кости.

Что касается математической теории, то она попросту не нуждается в том, чтобы распределение возможных исходов было равномерным. Несколько огрубляя положение вещей, можно сказать, что математическая теория вероятностей, вытекающая логически из определенных аксиом, учит тому, как по известным распределениям одних случайных событий или величин находить распределения различным образом связанных с ними других случайных событий или величин. Например, по аксиоме сложения вероятность выпадения четного числа очков есть

$$p_{\text{чет}} = p_2 + p_4 + p_6,$$

независимо от того, одинаковы ли и, вообще, чему равны вероятности  $p_2$ ,  $p_4$  и  $p_6$  выпадения грани с числом очков 2, 4 и 6.

Откуда известны исходные распределения — это вопрос, лежащий за пределами математики. Алгебра, например, учит, как для квадратного уравнения вычислить корни  $x_1$  и  $x_2$ , т. е. как выразить  $x_1$  и  $x_2$  через известные коэффициенты уравнения  $a$ ,  $b$  и  $c$ . Откуда мы знаем коэффициенты — этим алгебра не интересуется. В сущности, именно внематематический вопрос об источнике исходных распределений и был корнем различных воззрений на вероятность и теорию вероятностей до ее аксиоматизации, т. е. до того, как она раз и навсегда была освобождена от обязанности отвечать на этот вопрос. Ответ на него дает измерение, т. е. статистический опыт.

## § 2. Законы распределения случайных величин

Из сказанного ясно, почему физик (как и любой другой «потребитель» математической теории вероятностей), говоря о вероятности, всегда имеет в виду относительную частоту в статистическом ансамбле, в достаточно обширном статистическом

коллективе. Именно так он интерпретирует и закон распределения вероятностей случайной величины  $\xi$ , т. е. числовой характеристики исхода испытаний. Если возможные значения  $x$  величины  $\xi$  лежат в интервале  $a < x \leq b$  (в частности, в неограниченном интервале  $-\infty < x < \infty$ ), то закон распределения задается *интегральной функцией* распределения

$$W(x) = P\{a < \xi \leq x\}$$

— неотрицательной неубывающей функцией  $x$ , дающей вероятность попадания  $\xi$  в субинтервал  $a < \xi \leq x$  (случайное событие). При этом  $W(a) = 0$ , а по аксиоме II  $W(b) = P\{a < \xi \leq b\} = 1$ .

Если с ростом  $x$  функция  $W(x)$  возрастает скачками высоты  $p_i$  в точках  $x = x_i$ , а между этими точками постоянна (ступенчатая функция), то мы имеем частный случай дискретных возможных значений  $\xi$ , причем  $p_i$  — это вероятность события  $\xi = x_i$ :

$$p_i = P\{\xi = x_i\}.$$

Числа  $p_i$  аналогичны интенсивностям дискретных линий в линейчатом спектре или массам, сосредоточенным в точках  $x_i$  на оси  $x$ , причем полная интенсивность или полная масса равна единице:

$$\sum_i p_i = 1.$$

Если же функция распределения  $W(x)$  дифференцируема, т. е. существует *плотность вероятности*  $w(x) = dW(x)/dx$ , то это означает, что возможна непрерывная совокупность значений  $x$ , причем

$$w(x) dx = P\{x < \xi \leq x + dx\}, \quad \int_a^b w(x) dx = W(b) = 1.$$

Неотрицательная функция  $w(x)$  аналогична плотности интенсивности в сплошном спектре или плотности массы, непрерывно распределенной по интервалу  $(a, b]$  оси  $x$ .

Как сказано, для физика числа  $p_i$  или  $w(x)dx$  — это относительные частоты в статистическом ансамбле, т. е. доля тех реализаций случайного явления в обширном статистическом опыте, которые привели к значениям  $\xi = x_i$  при дискретных возможных значениях или к попаданию  $\xi$  в интервал  $(x, x + dx]$  при непрерывных возможных значениях.

Пусть, например, капелька краски внесена в момент  $t = 0$  в сосуд с водой, причем соблюдены необходимые предосторожности для того, чтобы не возникло движения воды. Тогда дальнейшее поведение капельки будет определяться диффузией

краски. Концентрацию краски  $c(\mathbf{r}, t)$  в любой точке  $\mathbf{r}$  в любой момент  $t > 0$  можно однозначно подсчитать, решив уравнение диффузии с заданными начальными и граничными условиями. Таким образом, это вполне детерминированное явление, подчиненное определенной динамической закономерности. Но если говорить о поведении отдельных частиц или молекул краски, то перед нами статистический опыт. Даже малая капля краски, скажем, объемом в  $10^{-3}$  мм<sup>3</sup>, содержит более  $10^{15}$  молекул. Все они находятся в практически тождественных условиях и совершают в воде хаотическое броуновское движение, которое и приводит к диффузионному расплыванию капли. Следовательно, здесь реализуется весьма обширный ансамбль молекул краски и ее количество  $c(\mathbf{r}, t)d^3r$  в элементе объема  $d^3r = dx dy dz$  в момент  $t$  — это те молекулы, которые случайно оказались к моменту  $t$  в этом элементе объема. Относительная концентрация  $c(\mathbf{r}, t)$  (т. е. нормированная так, что по всему объему сосуда  $\int c(\mathbf{r}, t)d^3r = 1$ ) — это для физика и есть плотность вероятности попадания отдельной молекулы в  $d^3r$  к моменту  $t$ , если движение началось из точки  $\mathbf{r} = 0$  в момент  $t = 0$ . Детерминированное поведение концентрации  $c(\mathbf{r}, t)$  — это проявление устойчивости относительных частот в достаточно обширном ( $10^{15}$ !) ансамбле.

В общем случае распределение вероятностей может быть смешанным — дискретно-непрерывным, — подобно суперпозиции линейчатого и сплошного спектров. Для лаконичной записи, охватывающей единым образом как дискретную, так и непрерывную компоненты, удобно пользоваться интегралом Стилтеса по  $dW(x)$ , где приращение  $dW(x)$  функции распределения на интервале  $(x, x + dx]$  может быть как бесконечно малым (сплошной спектр,  $dW(x) = w(x)dx$ ), так и конечным (дискретный спектр,  $dW(x) = W(x+0) - W(x) = p$ ). С помощью математического понятия интеграла Стилтеса можно, например, записать среднее значение какой-либо детерминированной функции  $f$  от случайной величины  $\xi$  в виде<sup>1)</sup>

$$\langle f(\xi) \rangle = \int_a^b f(x) dW(x), \quad (2.1)$$

независимо от того, дифференцируема функция  $W(x)$  или нет. Физики, однако, предпочитают иную, хотя и менее «строгую»,

<sup>1)</sup> Статистическое усреднение, т. е. усреднение по распределению вероятностей (или, как скажет физик, усреднение по ансамблю), мы будем обозначать угловыми скобками или, когда это будет удобно, просто чертой над усредняемой величиной.

но наглядную трактовку. Считается, что плотность  $w(x)$  имеет смысл везде и, соответственно,

$$\langle f(\xi) \rangle = \int_a^b f(x) w(x) dx. \quad (2.2)$$

В дискретном случае  $w(x)$  переходит в сумму дельта-функций в точках  $x = x_i$  с весами  $p_i$ :

$$w(x) = \sum_i p_i \delta(x - x_i). \quad (2.3)$$

Интеграл (2.2) тотчас приводит тогда к «дискретной» формуле

$$\langle f(\xi) \rangle = \sum_i f(x_i) p_i.$$

Представление (2.3) отвечает не только строго дискретным возможным значениям, но и непрерывным, при условии, что плотность вероятности  $w(x)$  имеет *достаточно острые пики*. Ширина этих пиков должна быть мала по сравнению с масштабами, фигурирующими в данной конкретной задаче. Следует, однако, иметь в виду, что идеализация этих пиков в виде дельта-выбросов, принятая в начале расчета, не всегда приводит к тому же результату, какой получается при переходе к пределу (2.3) уже в окончательных выражениях. С этой оговоркой неслучайную (детерминированную) величину  $\xi$ , достоверно принимающую значение  $x_0$ , можно почти всегда рассматривать как случайную величину с предельно острой плотностью вероятности  $w(x) = \delta(x - x_0)$ . Всякая динамическая теория предстает тогда перед нами как частный случай статистической теории, в которой сделан переход к предельно острым распределениям для величин, с которыми она оперирует. В подобном взгляде нет ничего парадоксального, в особенности после того, как волновая механика уже давно показала, что наиболее фундаментальные законы природы имеют статистический характер, а высказанный в свое время тезис, будто бы «наука — враг случайности», потерпел полный провал.

Практическая трактовка вероятности, как относительной частоты в достаточно обширном ансамбле, охватывает, разумеется, и многомерные случайные величины, т. е. совокупности случайных величин. В приведенном примере диффузии краски случайным был радиус-вектор каждой ее молекулы, т. е. совокупность трех случайных координат молекулы. Более того, в этом примере мы, по сути дела, рассматривали не случайные величины, а случайную функцию: плотность вероятности (концентрация)  $c$  зависела и от времени  $t$ . Такого рода обобщение классической



теории вероятностей будет в дальнейшем находиться в центре нашего внимания. Сейчас важно подчеркнуть лишь то, что и это обобщение не уведет от частотной трактовки законов распределения вероятностей.

Мы обратимся теперь к одному специальному дискретному закону распределения — биномиальному закону, — который не только позволит познакомиться с некоторыми интересными фактами и приложениями, но и послужит для дополнительной подготовки к переходу от случайных величин к случайным функциям.

### § 3. Биномиальный закон распределения

Если выполнение *конечного* числа испытаний можно расположить определенным образом во времени, например, приурочив испытание номера  $\nu$  к моменту времени  $t_\nu = \nu t$  (или как-либо иначе), то мы получим тем самым некоторую *модель случайного процесса*, для рассмотрения которой достаточны представления и методы классической теории вероятностей. В качестве примера рассмотрим хорошо известную задачу теории вероятностей — задачу Бернулли (1713 г.), в которой речь идет о последовательности *независимых* испытаний и которая может служить одной из наиболее ярких иллюстраций того, как одна и та же математическая схема (динамическая или статистическая) охватывает множество совершенно разнородных явлений.

Перечислим ряд конкретных вопросов, непосредственно сводящихся к математической схеме задачи Бернулли. Для части этих вопросов не требуется какой-либо локализации испытаний во времени, другие же допускают или даже предполагают такую локализацию, т. е. касаются процессов, разворачивающихся со временем.

1. Урновая задача. В урне лежат белые и черные шары, причем вероятность вынимания белого шара (событие  $A$ ) есть  $P(A) = p$  и соответственно для черного шара (событие  $\bar{A}$ )  $P(\bar{A}) = q = 1 - p$ . Производится  $N$  испытаний, причем каждый раз вынутый шар кладется обратно и шары перемешиваются (этим обеспечивается независимость испытаний). Какова вероятность  $P_N(n)$  того, что при  $N$  испытаниях *какие-либо*  $n$  раз будет вынут белый шар?

2. Бросание монеты. Какова вероятность того, что при  $N$  бросаниях  $n$  раз выпадет «орел», если вероятность выпадения «орла» при одном бросании есть  $p = 1/2$ ? Или какова вероятность того, что из  $N$  новорожденных  $n$  окажутся мальчиками, если вероятность рождения мальчика  $p = 0,51$ ?

3. Флуктуации плотности. В объеме  $V$  находится  $N$  молекул газа. Вероятность попадания молекулы в выделенный

из  $V$  объем  $v$  (событие  $A$ ) равна  $p$ . Какова вероятность того, что в  $v$  находится  $n$  каких-то молекул?

4. Флуктуации интенсивности. Складываются  $N$  колебаний вида  $a_i \cos \omega t$ , причем  $a_i = \pm a$  с вероятностями  $p$  и  $q$ . Какова вероятность того, что интенсивность суммарного колебания будет равна

$$J(N, n) = [n(+a) + (N - n)(-a)]^2 = (2n - N)^2 a^2,$$

т. е. амплитуда  $+a$  войдет  $n$  раз?

5. Телефонные вызовы. Телефонистка дежурит в течение времени  $T$ . Вероятность того, что за очень малый интервал  $\tau$  произойдет вызов, равна  $p$ . Какова вероятность  $n$  вызовов за все время дежурства?

6. Дробовой эффект в вакууме. Из накаленной нити вылетают электроны, причем вероятность вылета за весьма малое время  $\tau$  есть  $p$ . Какова вероятность того, что за время  $T$  на анод прилетит  $n$  электронов, т. е. поступит заряд  $ne$ ? Если в анодной цепи стоит интегрирующий прибор, суммирующий заряд за время  $T$ , то наш вопрос направлен к выяснению того, что покажет этот прибор в среднем и насколько он будет флуктуировать.

7. Случайные блуждания. Примем следующую модель движения брауновской частицы: она совершает скачки на отрезок  $a$  вправо или влево с вероятностями  $p$  и  $q$ . Какова вероятность того, что за  $N$  скачков частица уйдет на расстояние  $s(N, n) = (2n - N)a$ , т. е. сделает  $n$  шагов вправо? Если здесь  $p = q = 1/2$ , то можно сказать, что это одномерное движение «абсолютно пьяного человека»; если же  $A$  — шаг вправо,  $\bar{A}$  — шаг на месте, то это движение «нерешительного человека».

Три последних вопроса касаются процессов во времени. Они естественным образом связываются с представлением об испытаниях, производимых в последовательные моменты времени.

Общая постановка задачи, следовательно, такова: производится  $N$  независимых испытаний, при каждом происходит одно из противоположных событий  $A$  или  $\bar{A}$ . Какова вероятность того, что за  $N$  испытаний *какие-либо*  $n$  раз произойдет событие  $A$ ? Хотя решение задачи Бернулли хорошо известно, все же воспроизведем его здесь.

Итак,

$$P(A) = p, \quad P(\bar{A}) = q = 1 - p.$$

Какова вероятность того, что при  $N$  независимых испытаниях получится последовательность

$$A\bar{A}A\bar{A}\bar{A} \dots A,$$

т. е. на 1-м месте  $A$ , на 2-м месте  $\bar{A}$  и т. д., причем  $A$  произойдет всего  $n$  раз? Ответ очевиден:

$$pqrprq \dots p = p^n q^{N-n}.$$

Нас интересует событие  $B$ , состоящее в том, что  $A$  произошло  $n$  раз при *каких-то*  $n$  испытаниях из  $N$ . Появление  $n$  раз  $A$  при *определенных* испытаниях — это все частные случаи  $B$ , взаимно исключающие друг друга. По аксиоме сложения

$$P_N(n) = \sum_{\substack{\text{по всем частным} \\ \text{случаям } B}} p^n q^{N-n}.$$

Но число частных случаев  $B$  — это число способов, которыми можно выбрать  $n$  элементов из  $N$ , т. е.  $C_N^n = \frac{N!}{n!(N-n)!}$ . Следовательно,

$$P_N(n) = C_N^n p^n q^{N-n} \quad (3.1)$$

— закон распределения, называемый *биномиальным*, так как  $P_N(n)$  представляет собой  $n$ -й член разложения бинома  $(p+q)^N = \sum_{n=0}^N C_N^n p^n q^{N-n} = \sum_{n=0}^N P_N(n)$ . Написанная сумма должна быть равна единице, поскольку она выражает вероятность *какого-либо* из возможных значений  $n$ . Очевидно, это так и есть ввиду того, что  $p+q=1$ <sup>1)</sup>.

В более общем случае, когда при каждом испытании возможно наступление одного из  $k$  взаимно исключающих событий  $A_1, A_2, \dots, A_k$  с вероятностями  $p_1, p_2, \dots, p_k$  (разумеется,  $\sum_{i=1}^k p_i = 1$ ), вероятность того, что за  $N$  испытаний событие  $A_i$  произойдет  $n_i$  раз  $\left(\sum_{i=1}^k n_i = N\right)$ , будет выражаться соответствующим членом степенного разложения  $(p_1 + p_2 + \dots + p_k)^N$ , а именно:

$$P_N(n_1, n_2, \dots, n_k) = \frac{N!}{n_1! n_2! \dots n_k!} p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}.$$

<sup>1)</sup> Приведем решение так называемой задачи де Мере: каково должно быть число испытаний  $N$ , чтобы событие  $A$  произошло *хотя бы один раз*? Вероятность наступления  $A$  один, два, три и т. д. раз за  $N$  испытаний есть

$$P_N = \sum_{n=1}^N P_N(n) = \sum_{n=0}^N P_N(n) - P_N(0) = 1 - q^N = 1 - (1-p)^N.$$

С ростом  $N$  она стремится к единице. Принимая то или иное значение  $P_N$  в качестве «гарантирующей достоверность», нетрудно вычислить соответствующее  $N$ .

Рассмотрим некоторые свойства и следствия биномиального закона распределения.

Если обозначить через  $n_m$  то значение  $n$ , при котором вероятность  $P_N(n)$  максимальна, т. е. *наивероятнейшее* значение  $n$ , то из условия

$$P_N(n_m - 1) \leq P_N(n_m) \geq P_N(n_m + 1)$$

нетрудно получить, что

$$(N + 1)p - 1 \leq n_m \leq (N + 1)p.$$

Таким образом, имеется либо одно, либо два оптимальных значения. Если  $Np \gg 1$ , то

$$n_m \approx Np. \quad (3.2)$$

Найдем теперь *среднее* значение  $\bar{n}$  и *дисперсию*  $D[n] \equiv \sigma_n^2 \equiv \overline{(\Delta n)^2} \equiv \bar{n}^2 - \bar{n}^2$ . Для момента  $k$ -го порядка имеем

$$\begin{aligned} \bar{n}^k &= \sum_{n=0}^N n^k P_N(n) = \sum_{n=0}^N C_N^n n^k p^n q^{N-n} = \\ &= p \underbrace{\frac{\partial}{\partial p} p \frac{\partial}{\partial p} \dots}_{k \text{ раз}} \sum_{n=0}^N C_N^n p^n q^{N-n} = \left(p \frac{\partial}{\partial p}\right)^k (p+q)^N. \end{aligned}$$

После того как оператор  $p \frac{\partial}{\partial p}$  применен  $k$  раз, нужно учесть в полученном результате, что  $p+q=1$ . В частности,

$$\bar{n} = p \frac{\partial}{\partial p} (p+q)^N = pN (p+q)^{N-1} = pN, \quad (3.3)$$

$$\begin{aligned} \bar{n}^2 &= p \frac{\partial}{\partial p} p \frac{\partial}{\partial p} (p+q)^N = p \frac{\partial}{\partial p} pN (p+q)^{N-1} = \\ &= pN [(p+q)^{N-1} + p(N-1)(p+q)^{N-2}] = pN [1 + p(N-1)], \end{aligned} \quad (3.4)$$

и, следовательно,

$$\sigma_n^2 = \bar{n}^2 - \bar{n}^2 = Npq. \quad (3.5)$$

Заметим, что при больших  $N$  среднее значение  $\bar{n}$  совпадает, согласно (3.2), с наивероятнейшим. Что касается дисперсии  $\sigma_n^2$ , то она растет с числом испытаний  $N$  так же, как и среднее значение, т. е. пропорционально  $N$ . Поэтому среднее *относительное* отклонение

$$\frac{\sigma_n}{\bar{n}} = \sqrt{\frac{q}{pN}}$$

с ростом  $N$  убывает. Таким образом, с увеличением  $N$  флуктуации растут, но относительные флуктуации падают,

#### § 4. Примеры применения биномиального закона

Для уяснения физического смысла полученных результатов полезно обратиться к некоторым из поставленных выше конкретным задач.

Рассматривая газ в сосуде ( $N$  молекул в объеме  $V$ ), мы примем, что вероятность попадания молекулы в объем  $v$  есть  $p = v/V$ . Плотность газа в объеме  $v$  будет  $\rho_v = n/v$ . Следовательно,

$$\bar{\rho}_v = \frac{\bar{n}}{v} = \frac{pN}{v} = \frac{N}{V} = \bar{\rho} = \text{const},$$

т. е. в среднем плотность равномерна.

Дисперсия  $n$  определяет интенсивность флуктуаций плотности<sup>1)</sup>:

$$D[\rho_v] = D\left[\frac{n}{v}\right] = \frac{1}{v^2} D[n] = \frac{Npq}{v^2} = \frac{N}{vV} \left(1 - \frac{v}{V}\right),$$

или

$$D[\rho_v] = \overline{(\Delta\rho_v)^2} = \bar{\rho} \left(\frac{1}{v} - \frac{1}{V}\right).$$

Таким образом, абсолютные флуктуации плотности растут с уменьшением рассматриваемого объема  $v$  и с увеличением средней плотности, т. е. увеличением общего числа молекул  $N$  в объеме  $V$ . Обычно представляет интерес случай  $v \ll V$ , так что членом  $1/V$  можно пренебречь.

Что касается относительных флуктуаций, то

$$\frac{\overline{(\Delta\rho_v)^2}}{\bar{\rho}^2} = \frac{1}{\bar{\rho}} \left(\frac{1}{v} - \frac{1}{V}\right) \approx \frac{1}{\bar{\rho}v},$$

т. е. они тоже растут с уменьшением  $v$ , но падают с увеличением  $\bar{\rho}$ .

Отметим, в чем именно здесь заключается независимость испытаний. Мы приняли, что вероятность попадания молекулы в объем  $v$  равна  $p = v/V$  независимо от того, имеются ли и в каком количестве другие молекулы внутри или вне  $v$ . Другими словами, газ считается идеальным. При учете конечного объема молекул и их взаимодействия этого уже не будет: вероятность  $p$  при данном испытании (вероятность попадания *данной* молекулы в объем  $v$ ) будет зависеть от исходов остальных испытаний.

Рассмотрим теперь задачу о сложении колебаний. Биномиальный закон (3.1) выражает вероятность того, что  $n$  из  $N$  складываемых колебаний имеют амплитуду  $+a$ . Тем самым  $P_N(n)$

<sup>1)</sup> Как известно,  $D[ax + b] = a^2 D[x]$ .

есть вероятность значения результирующей интенсивности

$$J \equiv J(N, n) = (2n - N)^2 a^2.$$

Следовательно, средняя интенсивность будет

$$\bar{J} = \overline{(2n - N)^2} a^2 = (4\bar{n}^2 - 4N\bar{n} + N^2) a^2 = N^2 a^2 \left[ 1 + 4pq \left( \frac{1}{N} - 1 \right) \right].$$

Если  $p = q = 1/2$ , то  $J = Na^2$ , т. е. интенсивности складываются (полная некогерентность). При  $p = 1$  или  $q = 1$  получается  $\bar{J} = N^2 a^2$  (полная когерентность).

Дисперсия  $J$  есть

$$D[J] = \bar{J}^2 - \overline{J^2} = \overline{(2n - N)^4} a^4 - \bar{J}^2.$$

Таким образом, вычисление *флуктуаций интенсивности* — величины, квадратичной относительно  $n$ , — требует нахождения моментов  $n$  высших порядков, а именно  $\bar{n}^3$  и  $\bar{n}^4$ . Способ вычисления  $\bar{n}^k$  для распределения Бернулли уже был указан:  $k$ -кратное применение оператора  $p \frac{\partial}{\partial p}$  к биному  $(p + q)^N$ . Опуская промежуточные выкладки, приведем окончательный результат для случая полной некогерентности, когда  $p = q = 1/2$  и соответственно  $\bar{J} = Na^2$ . В этом случае дисперсия получается равной

$$D[J] = \overline{(\Delta J)^2} = 2N(N - 1) a^4,$$

так что относительная флуктуация интенсивности будет

$$\sqrt{\frac{\overline{(\Delta J)^2}}{\bar{J}^2}} = \sqrt{2 \left( 1 - \frac{1}{N} \right)}.$$

С ростом  $N$  относительная флуктуация вовсе не уменьшается, а, напротив, растет, приближаясь к значению  $\sqrt{2}$ . Таким образом, никакого «сглаживания» относительных флуктуаций здесь нет: при сложении двух колебаний относительная флуктуация интенсивности меньше, чем при сложении тысячи.

Привычное представление о «сглаживании» связано с моментом *второго* порядка. Если «шум» зависит от моментов высших порядков, то его роль может усиливаться с ростом  $N$ .

В этой связи следует остановиться на том, что называть «шумом» в том случае, когда прибор, например, квадратичен, т. е. измеряет не  $n$ , а  $n^2$ . Измеритель величины  $\xi = n^2$  показывает  $\bar{\xi} = \bar{n}^2$ . Он не обнаруживает никакого  $\xi$ -шума, хотя  $D[n] = \overline{n^2} - \bar{n}^2 \neq 0$ , т. е. само  $n$  флуктуирует. Шум для « $\xi$ -метра» определяется величиной  $D[\xi] = \overline{\xi^2} - \bar{\xi}^2 = \overline{n^4} - (\bar{n}^2)^2$  и связан, таким образом, с «флуктуацией флуктуации»  $n$ .

Отсюда видно, что причисление явления к флуктуационным или нефлуктуационным существенно зависит от того, какой величиной это явление характеризовать, — обстоятельство, подчеркнутое М. А. Леонтовичем ([25], § 34; см. также [26]). Пусть, например, речь идет об энергии  $U$  теплового излучения, заключенного в некотором объеме  $V$ . Она флуктуирует около своего среднего значения  $\bar{U}$ , так что дисперсия

$$D[U] = \bar{U}^2 - \bar{U}^2$$

есть мера интенсивности флуктуаций  $U$ . Но с точки зрения электродинамики энергия  $U$  равна интегралу от суммы квадратов напряженностей электрического и магнитного полей:

$$U = \frac{1}{8\pi} \int_V (E^2 + H^2) dV,$$

и, следовательно,

$$\bar{U} = \frac{1}{8\pi} \int_V (\bar{E}^2 + \bar{H}^2) dV.$$

Заметим теперь, что средние значения напряженностей в поле теплового излучения равны нулю ( $\bar{\mathbf{E}} = \bar{\mathbf{H}} = 0$ ), так что  $\bar{E}^2 = D[\mathbf{E}]$ ,  $\bar{H}^2 = D[\mathbf{H}]$  и

$$\bar{U} = \frac{1}{8\pi} \int_V (D[\mathbf{E}] + D[\mathbf{H}]) dV,$$

т. е. сама средняя энергия  $\bar{U}$  есть мера интенсивности флуктуаций  $\mathbf{E}$  и  $\mathbf{H}$ . Таким образом, мы имеем следующее расхождение терминологии при двух подходах:

1) *Энергетический подход*

Величина  $U$  со средним значением  $\bar{U}$ .

Флуктуация  $(\Delta U)^2$ .

2) *Электродинамический подход*

Величины  $\mathbf{E}$  и  $\mathbf{H}$  со средними значениями, равными нулю.

Флуктуация полей  $\bar{U}$ .

«Флуктуация флуктуации»  $(\Delta \bar{U})^2$ .

Обратимся теперь к задачам, в которых процесс «испытаний» разворачивается во времени, и в качестве первого примера возьмем описанную выше модель брауновского движения. Смещение частицы вправо за  $N$  шагов на величину

$$s = s(N, n) = (2n - N) a$$

имеет вероятность  $P_N(n)$ . Следовательно,

$$\bar{s} = (2\bar{n} - N) a = N(p - q) a,$$

т. е. в изотропном случае ( $p = q = 1/2$ ) среднее смещение будет  $\bar{s} = 0$ .

Разброс около  $\bar{s}$ , определяемый дисперсией, есть

$$D[s] = \overline{s^2} - \bar{s}^2 = D[2na - Na] = D[2na] = 4a^2 D[n] = 4a^2 Npq.$$

В изотропном случае, когда  $\bar{s} = 0$ , и  $D[s] = \overline{s^2}$ , получаем

$$\overline{s^2} = Na^2.$$

Если скачки происходят через равные промежутки времени  $\tau$ , так что  $t = N\tau$  — полное время, протекшее от начального момента, то

$$\overline{s^2} = \frac{a^2}{\tau} t.$$

Это — так называемая *диффузионная* зависимость пути от времени. В дальнейшем нам еще придется остановиться на связи рассматриваемой статистической схемы с диффузией. Мы увидим также, что простая рассмотренная модель представляет интерес и для радиофизики, будучи самым непосредственным образом связана с вопросом о флуктуациях фазы в автоколебательных системах.

## § 5. Дробовой эффект. Распределение Пуассона

Рассмотрим в заключение задачу о дробовом эффекте, вполне аналогичную задаче о телефонных вызовах.

Мы по-прежнему принимаем, что при каждом испытании вероятность события  $A$  (в данном случае это вылет электрона из катода в течение некоторого малого времени  $\tau$ ) имеет одно и то же значение  $p$ , не зависящее ни от исходов остальных испытаний, ни от номера рассматриваемого испытания, т. е. от эмиссии и присутствия других электронов и наличия объемного заряда. С учетом объемного заряда, вызывающего явление так называемой депрессии (подавления) дробового эффекта, задача существенно усложняется.

Перелет электрона от катода к аноду создает импульс анодного тока, длительность которого порядка времени пролета (в обычных электронных лампах — порядка  $10^{-9}$  сек), а интегральное значение равно заряду электрона  $e$ . Если средний анодный ток составляет, скажем,  $5$  ма, то каждую секунду на анод поступает примерно  $3 \cdot 10^{16}$  электронов. Отсюда ясно, что импульсы анодного тока густо перекрываются. Если бы они следовали вплотную друг за другом, но без перекрытия, то пролет указанного количества электронов занял бы около года. Но для нас сейчас нет необходимости рассматривать степень перекрытия импульсов, равно как и их форму. Под событием  $A$  мы будем понимать мгновенный акт вылета электрона из катода, т. е. *начало импульса* анодного тока.



Нас интересует заряд, поступивший на анод за некоторое время  $T = N\tau$ , где каждый интервал  $\tau$  — это одно из  $N$  испытаний, в результате которого констатируется либо наличие вылета в интервале  $\tau$  (с вероятностью  $p$ ), либо его отсутствие. Сразу же возникает вопрос о том, как быть с теми интервалами  $\tau$ , в которых оказалось два или более вылетов. Интуитивно ясно, что вероятность вылета  $p$  должна убывать с уменьшением  $\tau$ . Тогда при все более мелком дроблении промежутка времени  $T$  на части  $\tau = T/N$  вероятности двукратного, трехкратного и т. д. вылетов за время  $\tau$  будут уменьшаться, как величины соответственно второго, третьего и т. д. порядка малости относительно  $\tau$ . Мы можем поэтому рассчитывать на то, что при достаточно больших  $N$  (а никаких ограничений в этом отношении нет) для подавляющей доли интервалов  $\tau$  имеет место простая альтернатива: либо произошел один вылет (вероятность  $p$ ), либо интервал пустой (вероятность  $q = 1 - p$ ). Тем самым задача приводится к схеме Бернулли, и мы получаем, что вероятность  $n$  вылетов за время  $T$  (за  $N$  испытаний) есть  $P_N(n)$ . Сила тока в среднем за время  $T$  равна <sup>1)</sup>

$$I_T = \frac{ne}{T}.$$

Среднее значение тока  $I_T$  будет

$$\bar{I}_T = \frac{\bar{n}e}{T} = \frac{epN}{T} \equiv \bar{I}, \quad (5.1)$$

интенсивность флуктуаций  $I_T$  (дисперсия) —

$$D[I_T] \equiv \overline{(\Delta I_T)^2} = D\left[\frac{ne}{T}\right] = \frac{e^2}{T^2} D[n] = \frac{e^2 N p q}{T^2},$$

а относительное квадратичное отклонение —

$$\frac{\overline{(\Delta I_T)^2}}{\bar{I}_T^2} = \frac{q}{Np}.$$

Но в таком виде полученные результаты не могут претендовать на физическую значимость, так как они зависят от произвольного числа испытаний  $N$ .

Естественно, напрашивается предельный переход  $N \rightarrow \infty$ , осуществляемый при условии, что средняя сила тока  $\bar{I}$  фиксирована, т. е.

$$\bar{I} = n_1 e = \text{const},$$

<sup>1)</sup> В примере с плотностью газа мы полагали  $\rho_v = n/v$ , т. е. вводили величину, усредненную по рассматриваемой области  $v$ . Там это была область пространства, здесь — промежуток времени  $T$ .

где  $n_1$  — среднее число электронов, вылетающих за одну секунду. Согласно (5.1) это означает, что вероятность  $p$  вылета электрона за интервал времени  $\tau$  должна быть связана с  $N$  следующим образом:

$$p = \frac{n_1 T}{N} = n_1 \tau,$$

т. е.  $p$  должно стремиться к нулю вместе с  $\tau$ . Подставив  $p = n_1 T/N$  и  $q = 1 - n_1 T/N$  в выражение для  $D[I_T]$  и перейдя к пределу при  $N \rightarrow \infty$ , получаем теперь

$$\overline{(\Delta I_T)^2} = \frac{e \bar{I}}{T}, \quad \frac{\overline{(\Delta I_T)^2}}{\bar{I}^2} = \frac{e}{\bar{I} T}. \quad (5.2)$$

С ростом времени усреднения  $T$  флуктуации тока уменьшаются как абсолютно, так и относительно. С увеличением среднего тока  $\bar{I}$  происходит абсолютное усиление флуктуаций, но относительное их сглаживание. Формулы (5.2) — это основные формулы дробового эффекта без объемного заряда (без депрессии). Они явным образом отражают тот факт, что флуктуации тока — это следствие дискретности электрического заряда.

Дробовой эффект дает пример такой задачи, когда целесообразно считать, что число испытаний  $N$  за время  $T$  неограниченно растет, но при этом среднее число наступлений события  $A$  за единицу времени  $n_1$ , а значит и среднее их число за время  $T$ , т. е.  $\bar{n} = n_1 T$ , остается конечным. Тем самым вероятность  $A$  при отдельном испытании

$$p = \frac{n_1 T}{N} = \frac{\bar{n}}{N}$$

с ростом  $N$  стремится к нулю, т. е. в подавляющей доле испытаний событие  $A$  не наступает.

Во всех задачах такого типа (к ним относится и задача о телефонных вызовах, о сцинтилляциях, об ожидании трамвая, о соударениях молекул между собой или об их ударах о стенку и т. п.) напрашивается переход от общей формулы для  $P_N(n)$  к предельному выражению при  $N \rightarrow \infty$  и  $p \rightarrow 0$ , но фиксированном  $\bar{n}$ .

Заменяя в (3.1)  $p$  через  $\bar{n}/N$ , получаем

$$\begin{aligned} P_N(n) &= \frac{N(N-1) \dots [N-(n-1)]}{n!} \left(\frac{\bar{n}}{N}\right)^n \left(1 - \frac{\bar{n}}{N}\right)^{N-n} = \\ &= \frac{\bar{n}^n}{n!} \cdot 1 \cdot \left(1 - \frac{1}{N}\right) \dots \left(1 - \frac{n-1}{N}\right) \left(1 - \frac{\bar{n}}{N}\right)^{N-n}. \end{aligned}$$

В пределе при  $N \rightarrow \infty$  для вероятности  $n$  отсюда следует закон распределения Пуассона:

$$P(n) = \frac{\bar{n}^n}{n!} e^{-\bar{n}}. \quad (5.3)$$

Единственный параметр  $\bar{n}$ , входящий в этот закон, представляет собой именно среднее значение  $n$ , т. е.  $\sum_{n=0}^{\infty} nP(n) = \bar{n}$ . Все моменты высших порядков, конечно, тоже выражают через  $\bar{n}$  и, в частности,

$$\overline{n^2} = \bar{n}^2 + \bar{n}, \quad \overline{n^3} = \bar{n}^3 + 3\bar{n}^2 + \bar{n}.$$

Заметим, что если с самого начала считать вероятность наступления события  $A$  за время  $dt$  равной  $n_1 dt$  ( $dt$  вместо прежнего  $\tau$ ), то с точностью до величин второго порядка малости относительно  $dt$  совершенно строго получается закон Пуассона.

Легко видеть, что

$$\sum_{n=0}^{\infty} P(n) = 1,$$

как это и должно быть. Разумеется, среднее значение  $\bar{n}$  и стандарт  $\sigma_n$ , вычисленные при помощи распределения Пуассона, можно получить и посредством предельного перехода из выражений для этих величин, даваемых биномиальным законом, а именно:

$$\bar{n} = pN = n_1 T, \quad \sigma_n = \sqrt{Npq} = \sqrt{\bar{n}(1-p)} \rightarrow \sqrt{\bar{n}}.$$

С равномерным распределением вероятности события  $A$  ( $p = n_1 dt$ , где  $n_1 = \text{const}$ ) однозначно связан не только закон Пуассона, дающий вероятность того или иного числа  $n$  осуществлений события  $A$  за время  $T$ , но и вполне определенный (экспоненциальный) закон распределения для промежутка времени  $t$  от произвольно выбранного момента до *первого* наступления события  $A$ . Подойдем к этой задаче, отправляясь снова от схемы Бернулли.

Вероятность  $w(t)\Delta t$  осуществления двух независимых событий — того, что интервал  $(0, t)$  «пустой» и что в интервале  $(t, t + \Delta t)$  событие  $A$  произошло хотя бы один раз, — равна произведению вероятностей этих событий. Разобьем ось  $t$  на малые интервалы длины  $\tau$ , и пусть  $N$  таких интервалов содержится в  $(0, t)$ , так что  $t = N\tau$ , и  $\Delta N$  — в  $\Delta t$  ( $\Delta N = \frac{\Delta t}{\tau} = \frac{\Delta t}{t} N$ ). Вероятность «пустого» интервала  $(0, t)$  равна, следовательно,

$$q^N = (1 - p)^N = (1 - n_1 \tau)^N = \left(1 - \frac{n_1 t}{N}\right)^N,$$

а вероятность *хотя бы одного* наступления  $A$  в  $\Delta t$  равна (см. сноску на стр. 27)

$$1 - (1 - p)^{\Delta N} = 1 - \left(1 - \frac{n_1 t}{N}\right)^{\frac{\Delta t}{t} N}.$$

В пределе при  $N \rightarrow \infty$  первая вероятность переходит в  $e^{-n_1 t}$ , а вторая — в  $1 - e^{-n_1 \Delta t}$ , т. е. с точностью до первого порядка относительно  $\Delta t$  — в  $n_1 \Delta t$ . В результате получаем<sup>1)</sup>

$$w(t) = n_1 e^{-n_1 t}. \quad (5.4)$$

Согласно (5.4) среднее время ожидания есть

$$\bar{t} = \int_0^{\infty} t w(t) dt = \frac{1}{n_1},$$

что можно было написать сразу, исходя из определения  $n_1$ . Следует подчеркнуть, что момент, от которого отсчитывается время  $t$ , ничем не выделен. В частности, это может быть момент наступления события  $A$ , и тогда  $t$  будет промежутком времени между двумя осуществлениями  $A$  (соударениями молекул, телефонными вызовами, вылетами электронов и т. д.)<sup>2)</sup>.

## § 6. Предельная теорема Муавра — Лапласа

Другой предельный случай биномиального закона при  $N \rightarrow \infty$  получается тогда, когда  $p$  не стремится к нулю, а имеет определенное значение, отличное от нуля и единицы. Тем самым с ростом  $N$  растет также и  $\bar{n} = pN$ . Этот переход дается так называемой *локальной предельной теоремой*, доказанной Муавром (1730 г.) для альтернативы с  $p = q = 1/2$ , а затем Лапласом — для альтернативы с любыми  $p$  и  $q$  (конечно,  $p + q = 1$ ). Мы не будем рассматривать общий случай предельного перехода для  $P_N(n_1, n_2, \dots, n_k)$ , а ограничимся условиями теоремы Лапласа.

Формулируем первоначально какую-либо физическую задачу, для которой указанный предельный переход был бы естественным. Вполне подходящей для этого является уже рассмотренная модель брауновского движения. За  $N$  шагов частица

<sup>1)</sup> При проведенном предельном переходе мы перестаем следить за номерами испытаний, переносим внимание на фиксированный интервал времени  $\Delta t$ , в пределах которого число испытаний и их номера неограниченно растут вместе с  $N$ . Примером такого же изменения точки зрения, говоря гидродинамическим языком, с «материальной» на «локальную» может служить и часто используемый способ перехода от ряда к интегралу Фурье, когда вместо номеров (или частот) дискретных гармоник фиксируется частотный интервал  $\Delta\omega$ .

<sup>2)</sup> На первый взгляд это кажется странным. Кажется естественным, что если мы подошли к остановке трамвая как раз в тот момент, когда вагон уходит, то следующего трамвая придется ждать дольше, чем в том случае, когда мы приходим в промежуток между трамваями. Однако при отсутствии какого-либо расписания движения трамваев это неверно (см. [27], где рассмотрен вопрос о времени между соударениями молекул).

удаляется от начального положения на расстояние  $s(N, n) = (2n - N)a$  с вероятностью  $P_N(n)$ . Если шаг занимает время  $\tau$ , то расстояние  $s(N, n)$  проходится за время  $t = N\tau$ . Ввиду малости  $\tau$  реальным условиям наблюдения соответствуют промежутки времени  $t$ , охватывающие колоссальное количество  $N$  элементарных перемещений  $\pm a$ . Однако вероятности  $p$  и  $q = 1 - p$  перемещений  $+a$  и  $-a$  характеризуют рассматриваемую систему и не зависят от наблюдаемого числа шагов  $N$ . Таким образом, для всякого  $t \gg \tau$  здесь целесообразно пользоваться асимптотическим выражением  $P_N(n)$  при весьма больших  $N$ , но при фиксированном  $p$ .

Локальная теорема Муавра — Лапласа утверждает следующее:

*Если вероятность  $p$  события не зависит от номера испытания и отлична от 0 и 1, то вероятность  $P_N(n)$  дается при больших  $N$  асимптотическим выражением*

$$L_N(n) = \frac{e^{-x_n^2/2}}{\sqrt{2\pi Npq}}, \quad \text{где } x_n = \frac{n - pN}{\sqrt{Npq}}, \quad (6.1)$$

причем с ростом  $N$   $L_N(n)/P_N(n) \rightarrow 1$  равномерно для всех  $n$ , при которых  $x_n$  находится в конечном интервале, т. е.  $|x_n| < M$ , где  $M$  — некоторое положительное число.

Нетрудно убедиться, что распределение Лапласа (6.1) дает то же среднее значение  $\bar{n} = pN$  и дисперсию  $\sigma^2 = Npq$ , что и биномиальный закон. Можно поэтому записать  $L_N(n)$  в виде

$$L_N(n) = \frac{e^{-(n - \bar{n})^2/2\sigma^2}}{\sqrt{2\pi}\sigma}. \quad (6.2)$$

С ростом  $N$  происходит абсолютное расплывание, но относительное обострение распределения  $L_N(n)$ . Если фиксировать среднее значение и считать, что  $\sigma$  меняется независимо, то при  $\sigma \rightarrow 0$  распределение становится все острее, переходя в дельта-функцию.

Отметим, что аппроксимация распределения  $P_N(n)$  посредством  $L_N(n)$  при возрастании  $N$  становится весьма точной очень быстро (и тем быстрее, чем ближе  $p$  к  $1/2$ ). Если ввести  $\eta = P_N(n)/L_N(n)$ , то при  $p = 1/2$  имеем, например,

| $N$  | $n$ | $\eta$ |
|------|-----|--------|
| 25   | 15  | 1,0065 |
| 100  | 55  | 1,0030 |
| 1150 | 595 | 1,0001 |

На рис. 1 дан график, показывающий для частного случая  $n = \bar{n}$ , в каких областях плоскости  $(\bar{n}, 1/p)$  применимы полученные приближения. В области со штриховкой, наклоненной вправо,  $P_N(\bar{n})$  превышает  $P(\bar{n})$  не более чем на 1%, а в области со штриховкой, наклоненной влево,  $P_N(\bar{n})$  меньше, чем  $L_N(\bar{n})$ , тоже не более чем на 1%. Следует заметить, что вследствие выбора  $n = \bar{n}$  граничные кривые, строго говоря, имеют смысл только при целочисленных  $\bar{n}$ .

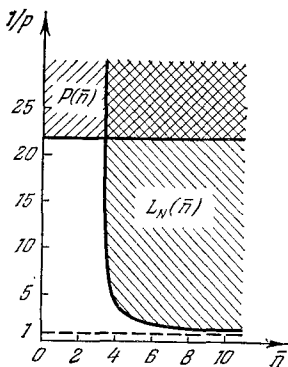


Рис. 1.

В верхнем правом углу графика пригодны обе асимптотические формулы, причем формула Лапласа удобней, так как она не содержит факториалов. В этой области — малых  $p$  и не слишком малых  $\bar{n}$  — распределение Лапласа тоже зависит только от одного параметра  $\bar{n}$ , поскольку при  $p \rightarrow 0$

$$\sigma^2 = Npq \rightarrow Np = \bar{n}.$$

Таким образом, распределение Лапласа может быть записано здесь в виде

$$L_N(n) = \frac{e^{-(n-\bar{n})^2/2\bar{n}}}{\sqrt{2\pi\bar{n}}}. \quad (6.3)$$

## § 7. Нормальный или гауссов закон распределения

В области применимости распределения Лапласа количество произведенных испытаний  $N$ , в сущности, уже не представляет интереса. Заранее известно, что за данное время  $t$  брауновская частица совершила чрезвычайно большое количество  $N$  (не существенно — какое именно) очень малых перемещений, и важно только распределение результирующего перемещения  $s$  (или результирующей амплитуды в задаче о сложении колебаний со случайными амплитудами  $\pm a$ ). Поэтому здесь также целесообразно перейти к «локальному» описанию<sup>1)</sup>, т. е. к непрерывному распределению  $w(s)ds$ , дающему вероятность попадания  $s$  в интервал  $(s, s + ds)$ .

Какова вероятность того, что  $x_n$  лежит в интервале  $(x, x + \Delta x)$ ? Этому интервалу  $x_n$  отвечает интервал  $n$  от  $n = \bar{n} + \sigma x$  до  $n = \bar{n} + \sigma(x + \Delta x)$ , т. е. в нем укладывается  $\sigma \Delta x = \sqrt{Npq} \Delta x$  значений  $n$ . Это число растет, как  $\sqrt{N}$ , в то

<sup>1)</sup> См. сноску <sup>1)</sup> на стр. 36.

время как относительная ширина интервала убывает, как

$$\frac{\sigma \Delta x}{\bar{n}} \sim \frac{1}{\sqrt{N}}.$$

Все значения  $L_N(n)$  для  $n$ , попадающих в интервал ширины  $\sigma \Delta x$ , практически одинаковы, а реализация любого из них является одним из взаимно исключающих частных случаев интересующего нас события. Поэтому, согласно аксиоме сложения, получаем

$$P\{x < x_n \leq x + \Delta x\} = \sum_{\bar{n} + \sigma x}^{\bar{n} + \sigma(x + \Delta x)} L_N(n) = L_N(n) \sigma \Delta x = \frac{e^{-x^2/2}}{\sqrt{2\pi}} \Delta x \quad (7.1)$$

— *нормальный или гауссов закон распределения.*

Теорему Муавра — Лапласа можно, следовательно, формулировать и так: биномиальный закон с  $0 < p < 1$  переходит при  $N \rightarrow \infty$  в нормальный закон для  $x$ , причем равномерно для любых конечных  $x$ .

Закон (7.1) получен для так называемой *нормированной случайной величины*

$$x_n = \frac{n - \bar{n}}{\sqrt{D[\bar{n}]}} ,$$

т. е. такой, у которой среднее значение равно нулю ( $\bar{x}_n = 0$ ), а дисперсия равна единице ( $D[x_n] = 1$ ). Если переписать гауссово распределение для ненормированной величины  $y$  со средним значением  $\bar{y}$  и с дисперсией  $D[y] = \sigma^2$ , то, поскольку  $x = (y - \bar{y})/\sigma$ , получим

$$w(y) dy = \frac{e^{-(y - \bar{y})^2/2\sigma^2}}{\sqrt{2\pi}\sigma} dy. \quad (7.2)$$

В частности, для перемещения  $s$  за время  $t$  при изотропном блуждании брауновской частицы мы нашли ранее (стр. 32), что  $\bar{s} = 0$ ,  $D[s] = \bar{s}^2 = Bt$ , где  $B = a^2/\tau$  — коэффициент диффузий. Следовательно,

$$w(s) ds = \frac{e^{-s^2/2Bt}}{\sqrt{2\pi Bt}} ds. \quad (7.3)$$

В начальный момент  $t = 0$  плотность  $w(s)$  превращается в  $\delta(s)$  (частица с достоверностью находится в точке  $s = 0$ ), а затем — с ростом  $t$  — происходит неограниченное расплывание распределения, как показано на рис. 2.

Нетрудно подсчитать, что для распределенной по нормальному закону (7.2) случайной величины  $y$  центральные моменты

имеют следующие значения:

$$\gamma_k = \overline{(y - \bar{y})^k} = \begin{cases} 0 & \text{при } k \text{ нечетном,} \\ 1 \cdot 3 \dots (k-1) \sigma^k & \text{при } k \text{ четном.} \end{cases}$$

Как уже было отмечено, распределение Гаусса имеет очень большое и очень общее значение. Задача Бернулли при больших  $N$  и  $\bar{n}$  — лишь одна из многих, приводящих к этому распределению. В дальнейшем мы еще обратимся к тем причинам, которые обуславливают особую роль и широкое распространение нормального закона. Здесь же достаточно напомнить следующее.

Вероятность состояния  $X$  ( $X$  — совокупность обобщенных координат  $q$  и импульсов  $p$ ) для системы в термостате с энергетической температурой  $\Theta = kT$  ( $k = 1,38 \cdot 10^{-16}$  эрг/град — постоянная Больцмана) дается так называемым каноническим распределением Гиббса:

$$w(X) dX = e^{\frac{\Psi - H(X)}{\Theta}} dX, \quad (7.4)$$

где  $\Psi$  — свободная энергия, а  $H(X)$  — гамильтониан системы. В тех случаях, когда  $H$  — неотрицательная билинейная форма, мы имеем именно гауссово (многочленное) распределение. Для

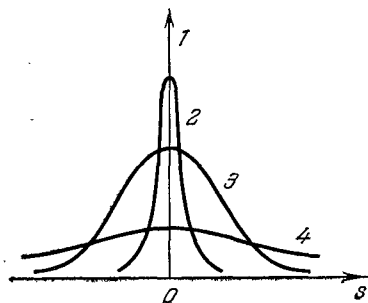


Рис. 2.

кинетической энергии, дающей распределение по скоростям, это справедливо всегда. Так, для одной частицы в силу того, что ее кинетическая энергия есть  $\frac{m}{2} (v_x^2 + v_y^2 + v_z^2)$ , из (7.4) следует нормальное распределение для  $v_x$ ,  $v_y$ ,  $v_z$ , т. е. максвелловское распределение:

$$w(v_x, v_y, v_z) dv_x dv_y dv_z = A e^{-\frac{m}{2\Theta} (v_x^2 + v_y^2 + v_z^2)} dv_x dv_y dv_z.$$

Отсюда для распределения абсолютной величины скорости  $v$  вытекает закон

$$w(v) dv = B e^{-\frac{mv^2}{2\Theta}} v^2 dv, \quad (7.5)$$

уже не являющийся нормальным.

Иногда вызывает недоумение следующий парадокс. Согласно (7.5) сколь угодно большие значения  $v$  имеют конечную вероятность

$$P\{v > v_0\} = B \int_{v_0}^{\infty} e^{-\frac{mv^2}{2\Theta}} v^2 dv.$$



Следовательно, конечную вероятность имеет и тот заведомо невозможный случай, когда одна молекула обладает кинетической энергией, превышающей энергию всего газа (всего термостата). Дело здесь как раз в предельном переходе от очень больших, но конечных  $N$  к  $N \rightarrow \infty$ . Ведь для  $n$  в биномиальном законе возможны значения только от 0 до  $N$ , распределения же Пуассона и Лапласа допускают любые  $n$ . При переходе от  $L_N(n)$  к непрерывному нормальному распределению верхняя граница отсутствует и для  $x$ .

Если рассматриваемая физическая система находится вблизи устойчивого состояния равновесия, совершая малые отклонения от него (флуктуации), то во многих случаях можно считать и потенциальную энергию  $U(q)$  неотрицательной квадратичной формой, т. е. распределение координат тоже гауссово.

Поясняя роль нормального распределения, мы сослались на примеры распределений Гиббса и Максвелла, т. е. на распределения, характеризующие *систему* случайных величин. Это естественное обобщение — переход от одной случайной величины к совокупностям таких величин и соответственно от одномерного распределения к многомерным — существенно как для непосредственных приложений, так и в качестве очередного шага на пути к теории случайных процессов. Достаточно заметить, что довольно часто используется способ задания случайной функции как *детерминированной* функции  $t$  и некоторой системы случайных параметров  $\alpha_1, \alpha_2, \dots, \alpha_n$ :

$$\xi(t) = f(t; \alpha_1, \alpha_2, \dots, \alpha_n).$$

Переход от одной случайной величины к системе таких величин особенно важен еще и потому, что именно здесь входит понятие о *статистической зависимости* между случайными величинами. Напомним, что означает статистическая независимость.

Пусть  $x_k$  — возможные значения случайной величины  $\xi_k$ . Величины  $\xi_1, \xi_2, \dots, \xi_n$  называются попарно независимыми (или, как еще говорят, независимыми в совокупности), если их совместная функция распределения распадается на произведение функций распределения каждой из них:

$$W(x_1, x_2, \dots, x_n) = W_{\xi_1}(x_1) \cdot W_{\xi_2}(x_2) \dots W_{\xi_n}(x_n).$$

Очевидно, такая же формула справедлива при этом и для плотностей вероятности:

$$w(x_1, x_2, \dots, x_n) = w_{\xi_1}(x_1) \cdot w_{\xi_2}(x_2) \dots w_{\xi_n}(x_n),$$

а любой смешанный момент порядка  $m = m_1 + m_2 + \dots + m_n$  распадается на произведение моментов:

$$\langle \xi_1^{m_1} \xi_2^{m_2} \dots \xi_n^{m_n} \rangle = \langle \xi_1^{m_1} \rangle \langle \xi_2^{m_2} \rangle \dots \langle \xi_n^{m_n} \rangle.$$

В частности, для двух независимых случайных величин  $\xi$  и  $\eta$  имеем

$$\overline{\xi\eta} = \bar{\xi} \cdot \bar{\eta}.$$

Однако обращение разности  $\overline{\xi\eta} - \bar{\xi} \cdot \bar{\eta}$  в нуль еще не означает статистической независимости  $\xi$  и  $\eta$ . Статистическая связь между  $\xi$  и  $\eta$ , мерой которой служит разность  $\overline{\xi\eta} - \bar{\xi} \cdot \bar{\eta}$ , называется *корреляцией* случайных величин  $\xi$  и  $\eta$ . Во многих случаях удобно пользоваться симметричной и безразмерной мерой корреляции — *коэффициентом корреляции*, который определен как

$$K_{\xi\eta} = \frac{\overline{\xi\eta} - \bar{\xi} \cdot \bar{\eta}}{\sqrt{D[\xi] \cdot D[\eta]}} = \frac{\overline{\xi\eta} - \bar{\xi} \cdot \bar{\eta}}{\sqrt{(\bar{\xi}^2 - \bar{\xi}^2)(\bar{\eta}^2 - \bar{\eta}^2)}}. \quad (7.6)$$

Нетрудно видеть, что при линейной связи  $\eta = a\xi + b$  имеем  $K_{\xi\eta} = 1$ . Легко также показать, что всегда  $|K_{\xi\eta}| \leq 1$ . При  $K_{\xi\eta} = 0$  говорят, что величины  $\xi$  и  $\eta$  некоррелированы, при  $K_{\xi\eta} < 0$  корреляция отрицательна, а случай  $K_{\xi\eta} = -1$  иногда (не очень удачно) называют антикорреляцией. Статистическая независимость случайных величин  $\xi$  и  $\eta$  влечет за собой их некоррелированность, но обратное в общем случае неверно.

Среднее значение суммы случайных величин всегда равно сумме их средних значений, но дисперсия суммы случайных величин есть

$$\begin{aligned} D\left[\sum_i \xi_i\right] &= \left\langle \left[\sum_i (\xi_i - \bar{\xi}_i)\right]^2 \right\rangle = \\ &= \sum_i D[\xi_i] + 2 \sum_{i < k} \sqrt{D[\xi_i] D[\xi_k]} K_{\xi_i \xi_k}. \end{aligned} \quad (7.7)$$

Таким образом, дисперсия суммы равна сумме дисперсий при достаточном условии попарной некоррелированности всех случайных величин (что, как сказано, еще не означает их попарной независимости).

Систему случайных величин  $\xi_1, \xi_2, \dots, \xi_n$  можно представлять себе как случайную точку или вектор  $\mathbf{p}$  в  $n$ -мерном пространстве. Поэтому такую систему часто называют, как известно,  $n$ -мерной случайной величиной.

Приведем в качестве примера общий вид нормального или гауссова закона распределения для  $n$ -мерной случайной величины  $\mathbf{p}$  (сама величина  $\mathbf{p}$  называется при этом нормальной или гауссовой). Общее выражение этого закона таково:

$$\omega(\mathbf{r}) d^n r = C e^{-H(\mathbf{r})} d^n r, \quad d^n r = dx_1 dx_2 \dots dx_n, \quad (7.8)$$

где  $H(\mathbf{r})$  — неотрицательная симметричная квадратичная форма:

$$H(\mathbf{r}) = \sum_{i,k} a_{ik} x_i x_k, \quad a_{ik} = a_{ki}, \quad (7.9)$$

постоянная  $C$  выражается через элементы матрицы  $A = (a_{ik})$  в соответствии с условием нормировки

$$\int_{-\infty}^{+\infty} \omega(r) d^n r = 1.$$

Таким образом, поверхности равной плотности вероятности — это семейство вложенных друг в друга подобных  $n$ -мерных эллипсоидов  $H(r) = \text{const}$ . Плотность вероятности максимальна в начале координат  $r = 0$ , где она равна  $C$ <sup>1)</sup>.

Обычно нормальное распределение записывается в канонической форме, т. е. через так называемую *корреляционную матрицу*  $B = (b_{ik})$ , где  $b_{ik}$  — моменты второго порядка:

$$b_{ik} = \overline{\xi_i \xi_k} = \int_{-\infty}^{+\infty} x_i x_k \omega(r) d^n r. \quad (7.10)$$

Если ввести стандарты  $\sigma_k$  компонент  $\xi_k$  рассматриваемой  $n$ -мерной случайной величины  $r$  ( $b_{ii} = \overline{\xi_i^2} = \sigma_i^2$ ), то можно записать  $b_{ik}$  через коэффициенты корреляции  $K_{ik}$ :

$$b_{ik} = \sigma_i \sigma_k K_{ik}.$$

Очевидно, матрица  $B$  симметрична, как и исходная матрица  $A$ . Мы имеем поэтому  $n(n+1)/2$  уравнений (7.10), выражающих все независимые  $b_{ik}$  через независимые  $a_{ik}$ . Если разрешить эти уравнения относительно  $a_{ik}$  и подставить полученные выражения  $a_{ik}$  через  $b_{ik}$  в (7.8) и (7.9), то распределение (7.8) принимает канонический вид:

$$\omega(r) d^n r = \frac{1}{(2\pi)^{n/2} |B|^{1/2}} \exp \left\{ -\frac{1}{2} \sum_{i,k} B_{ik}^{-1} x_i x_k \right\} d^n r. \quad (7.11)$$

Здесь  $|B| = \text{Det } B$ , а  $B_{ik}^{-1}$  — элементы матрицы  $B^{-1}$ , обратной матрице  $B$ .

<sup>1)</sup> Для распределения (7.8) средние значения  $\overline{\xi_i} = 0$ , как и вообще все моменты нечетных порядков. Если  $\overline{\xi_i} \neq 0$ , то это означает просто, что центр эллипсоидов смещен из начала координат в точку  $\bar{r} = \{\bar{\xi}_1, \bar{\xi}_2, \dots, \bar{\xi}_n\}$ . В этом случае в формуле (7.9)  $H$  будет квадратичной формой от компонент разности  $r - \bar{r}$ :

$$H(r) = \sum_{i,k} a_{ik} (x_i - \bar{\xi}_i) (x_k - \bar{\xi}_k).$$

В частности, для двумерной нормальной случайной величины закон распределения будет

$$\omega(x_1, x_2) dx_1 dx_2 = \frac{1}{2\pi \sqrt{b_{11}b_{22} - b_{12}^2}} \exp \left\{ -\frac{b_{22}x_1^2 + b_{11}x_2^2 - 2b_{12}x_1x_2}{2(b_{11}b_{22} - b_{12}^2)} \right\} dx_1 dx_2, \quad (7.12)$$

или, через стандарты и коэффициент корреляции  $K$  ( $b_{11} = \sigma_1^2$ ,  $b_{22} = \sigma_2^2$ ,  $b_{12} = b_{21} = \sigma_1 \sigma_2 K$ ),

$$\omega(x_1, x_2) dx_1 dx_2 = \frac{dx_1 dx_2}{2\pi \sigma_1 \sigma_2 \sqrt{1 - K^2}} \exp \left\{ -\frac{1}{2(1 - K^2)} \left( \frac{x_1^2}{\sigma_1^2} + \frac{x_2^2}{\sigma_2^2} - 2K \frac{x_1 x_2}{\sigma_1 \sigma_2} \right) \right\}. \quad (7.13)$$

В числе многих других особенностей нормального распределения следует указать и на ту, что для него некоррелированность влечет за собой статистическую независимость. Из двумерной формулы (7.13) это видно непосредственно, так как при  $K = 0$  плотность вероятности  $\omega(x_1, x_2)$  становится произведением двух одномерных нормальных плотностей. Но то же самое нетрудно показать и в общем  $n$ -мерном случае. Действительно, попарная некоррелированность  $\xi_i$  означает, что матрица  $B$  диагональна (отличны от нуля только  $b_{ii} = \sigma_i^2$ ). Следовательно, диагональна и обратная матрица  $B^{-1}$ , причем

$$B_{ii}^{-1} = \frac{1}{\sigma_i^2}, \quad |B|^{1/2} = \sigma_1 \sigma_2 \dots \sigma_n.$$

В результате получаем из (7.11)

$$\omega(\mathbf{r}) = \prod_i \frac{1}{\sqrt{2\pi} \sigma_i} e^{-x_i^2 / 2\sigma_i^2},$$

т. е.  $\omega(\mathbf{r})$  распадается на произведение  $n$  одномерных нормальных плотностей вероятности каждой из величин  $\xi_i$ .

Таким образом, для нормальных случайных величин некоррелированность и независимость равнозначны.

Заметим в заключение, что в физической литературе широко принято обозначать случайную величину той же буквой, что и ее возможные значения. Мы уже делали это выше, например, при рассмотрении модели брауновского движения: мы обозначали через  $n$  как случайное число шагов вправо, так и возможные значения  $n = 1, 2, \dots, N$  этого случайного числа. Такой прием сокращает количество употребляемых символов и в боль-

шинстве случаев не приводит к недоразумениям. В дальнейшем мы будем поступать так же, используя разные обозначения для случайной величины и ее возможных значений лишь тогда, когда без этого действительно возможно искажение смысла.

### Задачи

1. Написать совместную плотность вероятности случайных величин  $x = \sin \theta$  и  $y = \cos \theta$ , если известна плотность вероятности  $w(\theta)$  в интервале  $(-\pi, \pi)$ .

Решение. Запишем  $w_2(x, y)$  через условную плотность вероятности  $y$ :

$$w_2(x, y) = w_1(x) v(y|x).$$

Так как  $y$  — детерминированная функция  $x$ :

$$y = \begin{cases} +\sqrt{1-x^2} & \text{при } |\theta| < \pi/2, \\ -\sqrt{1-x^2} & \text{при } \pi/2 < |\theta| < \pi, \end{cases}$$

условная плотность есть

$$v(y|x) = \delta(y \mp \sqrt{1-x^2}).$$

Одномерное распределение  $x$  следует из заданного распределения  $\theta$ :

$$w(\theta) d\theta = w(\arcsin x) \frac{dx}{\sqrt{1-x^2}} = w_1(x) dx.$$

Таким образом,

$$w_2(x, y) = \frac{w(\arcsin x)}{\sqrt{1-x^2}} \delta(y \mp \sqrt{1-x^2})$$

с верхним знаком при  $|\arcsin x| < \pi/2$  и с нижним — при  $\pi/2 < |\arcsin x| < \pi$ . Разумеется, можно написать аналогичное выражение и через  $v(x|y)$ .

2. Какие распределения случайной величины  $x$  на интервале  $(-a, a)$  обеспечивают некоррелированность  $x$  и любой четной детерминированной функции  $f(x)$ ?

Решение. Достаточным условием некоррелированности  $x$  и  $f(x)$  является четность  $w(x)$ , так как при этом обращаются в нуль моменты  $\langle x f(x) \rangle$  и  $\langle x \rangle$ .

3. Каким условиям должна удовлетворять плотность вероятности случайной фазы  $\theta$  в интервале  $(-\pi, \pi)$  колебаний  $\xi = \sin(\omega t + \theta)$  и  $\eta = \cos(\omega t + \theta)$ , чтобы они были некоррелированы?

Решение. Коэффициент корреляции  $\xi$  и  $\eta$  пропорционален разности

$$\begin{aligned} \overline{\xi \eta} - \bar{\xi} \cdot \bar{\eta} &= [\langle \sin \theta \cos \theta \rangle - \langle \sin \theta \rangle \cdot \langle \cos \theta \rangle] \cos 2\omega t + \\ &+ \frac{1}{2} [\langle \cos^2 \theta \rangle - \langle \cos \theta \rangle^2 - \langle \sin^2 \theta \rangle + \langle \sin \theta \rangle^2] \sin 2\omega t. \end{aligned} \quad (1)$$

Представим плотность вероятности  $w(\theta)$  в виде

$$w(\theta) = \frac{1}{2\pi} + \sum_{k=1}^{\infty} (a_k \sin k\theta + b_k \cos k\theta).$$

Вычислив для такого распределения моменты, входящие в коэффициенты при  $\sin 2\omega t$  и  $\cos 2\omega t$  в соотношении (1), и приравняв эти коэффициенты нулю,

находим следующие два условия:

$$a_2 = 2\pi a_1 b_1, \quad b_2 = \pi (b_1^2 - a_1^2).$$

Если ввести амплитуды и фазы первой и второй гармоник  $w(\theta)$ :

$$a_i = A_i \sin \alpha_i, \quad b_i = A_i \cos \alpha_i, \quad i = 1, 2,$$

то полученные условия означают, что

$$A_2 = \pi A_1^2, \quad \alpha_2 = 2\alpha_1.$$

В частности, колебания  $\xi$  и  $\eta$  будут некоррелированы при отсутствии в  $w(\theta)$  первой и второй гармоник ( $A_1 = 0$ ,  $A_2 = 0$ ), а также, конечно, и при отсутствии всех гармоник, т. е. при равномерном распределении  $w(\theta) = 1/2\pi$ .

4. Найти дисперсию суммы случайного числа  $n$  независимых случайных величин  $x_i$ , обладающих одинаковыми дисперсиями  $\sigma^2$ .

Решение. Совместную функцию распределения  $W(y, n)$  случайных величин  $y = \sum_{i=1}^n x_i$  и  $n$  можно записать через условную вероятность  $V(y|n)$ , умноженную на вероятность  $p_n$  значения  $n$ :

$$W(y, n) = p_n V(y|n).$$

Следовательно,

$$D[y] = \sum_n p_n D[y|n] = \sum_n p_n n \sigma^2 = \bar{n} \sigma^2.$$

5. Показать, что для  $n$  нормальных величин  $x_i$  всегда можно найти такие  $n$  линейных комбинаций  $y_k = \sum_i a_{ki} x_i$ , которые будут независимы в совокупности.

Решение. Справедливость данного утверждения следует из того, что всегда существует ортогональное преобразование, приводящее квадратичную форму  $H(\mathbf{r})$  к сумме квадратов. Геометрически это поворот осей в  $n$ -мерном пространстве  $x_i$ , при котором новые оси совпадают с главными осями эллипсоидов  $H = \text{const}$ .

6. Показать, что для нормальной величины  $x$  со средним значением  $\bar{x} = 0$  и дисперсией  $\overline{x^2} = \sigma^2$  для всякой детерминированной дифференцируемой функции  $f(x)$ , возрастающей при  $|x| \rightarrow \infty$  медленнее, чем  $e^{x^2/2\sigma^2}$ , справедлива формула

$$\langle x f(x) \rangle = \sigma^2 \langle f'(x) \rangle. \quad (1)$$

Решение. Интегрирование по частям дает

$$\begin{aligned} \langle x f(x) \rangle &= \int_{-\infty}^{+\infty} x f(x) e^{-x^2/2\sigma^2} \frac{dx}{\sqrt{2\pi} \sigma} = \\ &= -\sigma^2 \left\{ \frac{f(x)}{\sqrt{2\pi} \sigma} e^{-x^2/2\sigma^2} \Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} f'(x) e^{-x^2/2\sigma^2} \frac{dx}{\sqrt{2\pi} \sigma} \right\} = \sigma^2 \langle f'(x) \rangle. \end{aligned}$$

Обобщение этой формулы на функционалы от случайных гауссовых полей используется в теории случайных полей.

Если  $\bar{x} \neq 0$ , то формула (1) принимает вид

$$\langle xf(x) \rangle = \bar{x} \langle f(x) \rangle + \sigma^2 \langle f'(x) \rangle,$$

где  $\sigma^2 = \overline{x^2} - \bar{x}^2$ .

7. Пусть задана  $n$ -мерная случайная величина  $\xi = \{\xi_1, \dots, \xi_n\}$ , т. е. известна ее  $n$ -мерная плотность вероятности  $w_\xi(x_1, \dots, x_n)$ , где  $x_k$  — возможные значения  $\xi_k$ . Величина  $\eta = (\eta_1, \dots, \eta_n)$  связана с  $\xi$  функциональной зависимостью  $\eta = F(\xi)$ , где  $F$  — детерминированная однозначная вектор-функция с однозначной обратной функцией  $\xi = F^{-1}(\eta)$ . Выразить плотность вероятности  $w_\eta(y)$  через  $w_\xi(x)$ .

Решение. Однозначную связь  $\xi = F^{-1}(\eta)$  можно рассматривать как преобразование координат в  $n$ -мерном пространстве возможных значений:

$$x_i = F_i^{-1}(y_1, \dots, y_n), \quad i = 1, 2, \dots, n.$$

Как известно, элемент «объема»  $d^n x = dx_1 \dots dx_n$  преобразуется по формуле  $d^n x = Q(y) d^n y$ , где  $Q(y)$  — якобиан преобразования:

$$Q(y) = \text{Det} \left| \frac{\partial x_i}{\partial y_k} \right| = \text{Det} \left| \frac{\partial F_i^{-1}(y)}{\partial y_k} \right|.$$

В силу взаимной однозначности преобразования вероятность попадания  $\eta$  в область  $(y, y + dy)$  та же, что и  $\xi$  в область  $(x, x + dx)$ , т. е.

$$w_\eta(y) d^n y = w_\xi(x) d^n x.$$

Подставив в правую часть  $x = F^{-1}(y)$  и  $d^n x = Q(y) d^n y$ , находим

$$w_\eta(y) = w_\xi[F^{-1}(y)] \cdot Q(y).$$

## Г л а в а II

### СЛУЧАЙНЫЕ ИМПУЛЬСЫ

#### § 8. Постановка задачи

Мы обратимся теперь к одному частному виду случайных процессов — так называемым *пуассоновским импульсным процессам*. Этот практически важный класс процессов вместе с тем настолько специален, что для получения многих относящихся к нему результатов можно и здесь обойтись средствами классической теории вероятностей. Поясним, в чем здесь дело.

Мы уже отметили, что случайная функция времени  $t$  иногда может быть задана как детерминированная функция  $f$  от  $t$  и от некоторого *конечного* числа случайных величин  $\alpha_1, \dots, \alpha_n$ . Все статистические свойства функции  $\xi(t) = f(t; \alpha_1, \dots, \alpha_n)$  полностью определяются тогда  $n$ -мерной случайной величиной  $\alpha_1, \dots, \alpha_n$ , т. е. ее распределением  $\omega(\alpha_1, \dots, \alpha_n) d\alpha_1 \dots d\alpha_n$ .

Примером такого задания случайной функции могут служить тригонометрические суммы со случайными коэффициентами  $a_k, b_k$

$$\xi(t) = f[t; (a_k), (b_k)] = \sum_{k=0}^n (a_k \cos kt + b_k \sin kt),$$

определяющие случайные функции  $\xi(t)$ , периодические с периодом  $2\pi$ .

Пока число  $n$  случайных параметров конечно, мы можем оставаться в рамках тех понятий и методов, которыми оперирует классическая теория вероятностей, но переход к счетному множеству случайных параметров (в приведенном примере — к ряду Фурье) требует уже более общего подхода. Импульсные процессы можно, по крайней мере в простейших случаях, трактовать именно указанным способом, задавая их как детерминированные функции времени и конечной совокупности случайных параметров. Тем самым мы имеем возможность познакомиться



с этими процессами, не прибегая пока к общей теории случайных функций.

Рассмотрим процесс, представляющий собой суперпозицию одинаковых по форме импульсов. Пусть форма импульса описывается детерминированной функцией  $F(t)$ , причем импульсы могут различаться по величине или «амплитуде». Интересующий нас процесс запишется тогда в виде

$$\xi(t) = \sum_{\nu} a_{\nu} F(t - t_{\nu}), \quad (8.1)$$

где  $t_{\nu}$  — момент «возникновения»  $\nu$ -го импульса, а  $a_{\nu}$  — его «амплитуда». Мы называем  $t_{\nu}$  моментом возникновения лишь условно, т. е. не предполагаем, что  $F(t - t_{\nu}) = 0$  при  $t < t_{\nu}$ . Функция  $F(t)$  может и не обладать начальным значением  $t$ , до которого она равна нулю. Предполагается только, что  $F(t)$  достаточно быстро стремится к нулю при  $|t| \rightarrow \infty$ . Моменты времени  $t_{\nu}$  могут быть связаны с любой характерной точкой — с каким-либо из экстремумов  $F(t)$  или с какой-нибудь из точек перехода через нуль, если таковые имеются, и т. п.

Если параметры  $a_{\nu}$ ,  $t_{\nu}$  — случайные величины, то  $\xi(t)$  — случайная функция времени: при всяком  $t$  значение  $\xi(t)$  само будет случайной величиной. Таким образом, мы имеем здесь пример описанного способа задания случайной функции как детерминированной функции некоторой совокупности случайных параметров (в данном случае  $a_{\nu}$  и  $t_{\nu}$ ).

Укажем на некоторые физические вопросы, приводящие к рассмотрению случайных процессов вида (8.1). Таким процессом является, например, ток или напряжение на выходе четырех-полюсника, на вход которого воздействует аналогичный процесс, но состоящий из очень коротких толчков — в пределе дельта-импульсов:

$$I_{\text{вх}}(t) = \sum_{\nu} a_{\nu} \delta(t - t_{\nu}). \quad (8.2)$$

Под «очень короткими» понимаются при этом толчки, длительность которых гораздо меньше временных постоянных четырех-полюсника. В частности, если этому условию удовлетворяет время пролета электрона в электронной лампе, то анодный ток можно записать в виде (8.2) с  $a_{\nu} = e = \text{const}$ . Если же необходимо учитывать время пролета и тем самым форму импульсов анодного тока, вызванных отдельными электронами, то тот же анодный ток можно записать в виде (8.1), по-прежнему с  $a_{\nu} = e$ , если

$$\int_{-\infty}^{+\infty} F(t) dt = 1.$$

Другой пример возникновения случайного процесса вида (8.1) — это отражение волнового импульса, посланного в среду, содержащую хаотически вкрапленные неоднородности со случайными коэффициентами отражения или рассеяния. Отраженные от неоднородностей импульсы будут иметь в отсутствие дисперсии ту же форму, что и первичный импульс, но вернуться со случайными запаздываниями и амплитудами.

Возможна еще более общая постановка задачи, которую далее мы и будем рассматривать. Она состоит в том, что случайной предполагается форма самих импульсов, а именно — импульс описывается детерминированной функцией  $F$  от  $t - t_v$  и от совокупности некоторого конечного числа  $m$  случайных параметров, т. е. от  $m$ -мерной случайной величины  $\mathbf{a}_v = \{a_{v1}, \dots, a_{vm}\}$ :

$$\xi(t) = \sum_v F(t - t_v, \mathbf{a}_v). \quad (8.3)$$

Здесь мы снова используем, таким образом, тот же ограниченный способ задания случайной функции  $F$  как детерминированной функции конечного числа случайных параметров. При этом компоненты  $\mathbf{a}_v$  могут быть и статистически зависимыми между собой.

Относительно вероятностных свойств случайных параметров  $\mathbf{a}_v$  и  $t_v$ , определяющих статистические характеристики процесса  $\xi(t)$ , мы сделаем следующие простейшие предположения:

1. Все  $\mathbf{a}_v$  и  $t_v$  статистически независимы между собой, и их распределения не зависят от номера импульса  $v$ . Тогда для системы, состоящей из любого количества  $n(m+1)$  параметров  $\mathbf{a}_v$  и  $t_v$ , достаточно задать только функции распределения  $w_a(\mathbf{a}) d\mathbf{a}$  и  $w_t(t) dt$ , поскольку функция распределения системы распадается на множители:

$$\begin{aligned} w(\mathbf{a}_1, \dots, \mathbf{a}_n; t_1, \dots, t_n) d\mathbf{a}_1 \dots d\mathbf{a}_n dt_1 \dots dt_n = \\ = \prod_{v=1}^n w_a(\mathbf{a}_v) d\mathbf{a}_v w_t(t_v) dt_v. \end{aligned} \quad (8.4a)$$

Здесь  $d\mathbf{a}_v \equiv d^m a_v = da_{v1} \dots da_{vm}$ .

2. Вероятность появления импульса в промежутке времени от  $t$  до  $t + dt$  не зависит от  $t$  и пропорциональна  $dt$ , т. е.

$$w_t(t) dt = n_1 dt, \quad n_1 = \text{const.} \quad (8.4b)$$

Как уже отмечалось (§ 5), это предположение означает, что вероятность появления  $n$  импульсов в интервале времени  $T$  дается распределением Пуассона:

$$P(n) = \frac{\bar{n}^n e^{-\bar{n}}}{n!}, \quad \bar{n} = n_1 T. \quad (8.5)$$

В этих предположениях мы вычислим теперь функцию распределения  $\xi(t)$ .

Чтобы избежать оперирования с бесконечной совокупностью случайных параметров  $a_n, t_n$ , мы поступим следующим образом. Выделим интервал времени  $(-T/2, T/2)$  настолько большой по сравнению с длительностью отдельного импульса и со средним временем  $1/n_1$  между импульсами, чтобы можно было пренебречь краевыми эффектами, т. е. чтобы учет (или неучет) тех импульсов, которые частично срезаются краями интервала  $(-T/2, T/2)$ , не играл роли. Выполнимость этого последнего требования как раз и связана с условием достаточно быстрого стремления  $F(t)$  к нулю при  $|t| \rightarrow \infty$ . Вопрос о том, насколько быстрым должно быть уменьшение  $F(t)$ , мы пока отложим, отметив лишь, что поставленное условие пренебрежимости краевыми эффектами может быть выполнено даже тогда, когда интервал, в котором  $F(t) \neq 0$ , не является конечным.

Мы хотим найти распределение случайной величины  $\xi(t)$  (а при всяком фиксированном  $t$  это случайная величина), т. е. вероятность события  $B$ , состоящего в том, что  $\xi(t)$  находится в интервале  $(x, x + dx)$ :

$$P\{B\} = P\{x < \xi(t) \leq x + dx\} = w(x) dx.$$

Событие  $B$  может реализоваться при осуществлении какого-либо из событий  $A_0, A_1, \dots, A_n, \dots$ , где  $A_n$  состоит в том, что в интервале  $(-T/2, T/2)$  появилось  $n$  импульсов. События  $A_n$  несовместимы, так что по формуле полной вероятности

$$P\{B\} = \sum_{n=0}^{\infty} P\{A_n\} P\{B|A_n\}.$$

Очевидно,  $P\{A_n\}$  есть не что иное, как вероятность  $P(n)$ , даваемая законом Пуассона. Введем теперь плотность условной вероятности  $P\{B|A_n\}$ , т. е. вероятности события  $B = \{x < \xi(t) \leq x + dx\}$  при условии, что в  $(-T/2, T/2)$  возникло  $n$  импульсов:

$$P\{B|A_n\} = v(x|n) dx. \quad (8.6)$$

Тогда

$$w(x) dx = \sum_{n=0}^{\infty} P(n) v(x|n) dx \quad (8.7)$$

и, следовательно, среднее значение какой-либо функции  $f(\xi)$  может быть вычислено в два этапа:

$$\begin{aligned} \overline{f(\xi)} &= \int_{-\infty}^{+\infty} f(x) w(x) dx = \\ &= \sum_{n=0}^{\infty} P(n) \int_{-\infty}^{+\infty} f(x) v(x|n) dx = \sum_{n=0}^{\infty} P(n) \overline{f(\xi)}^{(n)}, \end{aligned} \quad (8.8)$$

где  ${}^n\overline{f}(\xi)$  — условное среднее [при условии, что в  $(-T/2, T/2)$  было  $n$  импульсов].

Итак, согласно (8.7), для нахождения  $w(x)$  надо располагать условной плотностью вероятности  $v(x|n)$ . Можно сказать, что  $v(x|n)$  — это плотность вероятности случайной функции

$$\xi_n(t) = \sum_{v=1}^n F(t - t_v, a_v), \quad (8.9)$$

зависящей от конечного числа параметров  $a_v$  и  $t_v$ , относящихся к тем  $n$  импульсам, которые по условию появились в интервале  $(-T/2, T/2)$ . Для нахождения  $v(x|n)$  целесообразно воспользоваться очень сильным аппаратом так называемых *характеристических функций*.

### § 9. Характеристическая функция

Характеристические функции играют большую роль как в доказательствах многих важных теорем теории вероятностей, так и при решении конкретных задач. Характеристическая функция *однозначно* связана с функцией распределения  $W(x)$  и имеет по сравнению с  $W(x)$  очень существенное преимущество. При сложении *независимых* случайных величин, когда распределение суммы вычисляется по формуле свертки (композиция или символическое перемножение функций распределения), характеристическая функция суммы получается простым перемножением характеристических функций слагаемых. Это особенно важно для центральной предельной теоремы теории вероятностей, где речь идет именно о распределении суммы *независимых* случайных величин (§ 13).

Характеристическая функция случайной величины  $x$  (как уже было сказано, мы будем сплошь и рядом обозначать случайную величину и принимаемые ею значения одной и той же буквой) есть среднее значение (математическое ожидание) величины  $e^{iux}$ , где  $u$  — вещественный параметр:

$$\varphi_x(u) = \langle e^{iux} \rangle = \int_{-\infty}^{+\infty} e^{iux} dW(x) \quad (9.1)$$

(разумеется, здесь принято, что  $\overline{a+ib} = \bar{a} + i\bar{b}$ ). Интеграл Фурье — Стильеса (9.1) всегда сходится, так как

$$|\varphi_x(u)| \leq \int_{-\infty}^{+\infty} |e^{iux}| dW(x) = 1.$$

Нетрудно видеть, далее, что всегда  $\varphi_x(0) = 1$  и что  $\varphi_x(-u) = \varphi_x^*(u)$ . Приведем два примера характеристической функции.

Для *нормального распределения*

$$dW(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-a)^2/2\sigma^2} dx$$

имеем

$$\varphi_x(u) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} e^{iux} e^{-(x-a)^2/2\sigma^2} dx = e^{iau - \sigma^2 u^2/2}.$$

Чем острее распределение (чем меньше  $\sigma$ ), тем шире характеристическая функция. Конечно, это не особенность нормального распределения, а известное общее свойство пары функций, связанных интегральным преобразованием Фурье. Если  $\sigma \rightarrow 0$ , т. е.  $w(x) \rightarrow \delta(x-a)$ , то  $\varphi_x(u) \rightarrow e^{iau}$  (в частности, при  $a=0$  имеем  $\varphi_x(u) \rightarrow 1$ ).

Для *распределения Пуассона*

$$P(n) = \frac{\bar{n}^n}{n!} e^{-\bar{n}}$$

получаем

$$\varphi_n(u) = \sum_{n=0}^{\infty} e^{iun} P(n) = e^{-\bar{n}} \sum_{n=0}^{\infty} \frac{(\bar{n}e^{iu})^n}{n!} = \exp\{\bar{n}(e^{iu} - 1)\}.$$

Если для какого-либо  $k > 0$  конечен абсолютный момент  $\beta_k = |\overline{x^k}|^k$ , то выражение  $\varphi_x(u)$  можно  $k$  раз продифференцировать по  $u$ . Поскольку

$$e^{iux} = 1 + \frac{iux}{1!} + \dots + \frac{(iux)^k}{k!} + R_k(iux),$$

получаем

$$\varphi_x(u) = \sum_{m=0}^k \frac{\alpha_m (iu)^m}{m!} + o(u^k), \quad \alpha_m = \overline{x^m}. \quad (9.2)$$

Если все  $\beta_k$  конечны, то имеет смысл разложение в бесконечный ряд

$$\varphi_x(u) = \sum_{m=0}^{\infty} \frac{\alpha_m (iu)^m}{m!}. \quad (9.3)$$

Заметим, что  $m$ -й момент  $\alpha_m$ , если он конечен, вычисляется при помощи характеристической функции, как значение ее  $m$ -й производной по  $iu$  в нуле:

$$\alpha_m = \frac{\partial^m \varphi_x(u)}{\partial (iu)^m} \Big|_{u=0}. \quad (9.4)$$

Таким образом, нахождение моментов, требующее — если оно производится при помощи функции распределения — вычисления интегралов, здесь осуществляется посредством дифференцирования, т. е., вообще говоря, более простой операции (правда, лишь для целых  $m$ ).

Вряд ли надо пояснять, что для многомерной случайной величины  $\mathbf{r} = \{x_1, x_2, \dots, x_n\}$  характеристическая функция определяется точно таким же образом, т. е. как функция, сопряженная по Фурье:

$$\varphi_{\mathbf{r}}(\mathbf{u}) = \int_{-\infty}^{+\infty} e^{i\mathbf{u}\mathbf{r}} dW(\mathbf{r}), \quad (9.5)$$

где  $\mathbf{u}$  — вектор с компонентами  $u_1, u_2, \dots, u_n$ .

Для плотности вероятности величины  $z = x + y$ , где  $x$  и  $y$  — независимые случайные величины, имеем формулу композиции (свертки)

$$w(z) = w_1 * w_2 = \int_{-\infty}^{+\infty} w_1(x) w_2(z-x) dx = \int_{-\infty}^{+\infty} w_1(z-y) w_2(y) dy,$$

т. е. довольно сложную функциональную связь между  $w_1$ ,  $w_2$  и  $w$ . Для характеристической функции  $z$ , учитывая независимость  $x$  и  $y$ , получаем

$$\varphi_z(u) = \overline{e^{iuz}} = \overline{e^{iux}} \overline{e^{iuy}} = \varphi_x(u) \varphi_y(u).$$

Таким образом, композиции  $w$  или  $W$ , т. е. символическому перемножению функций распределения, соответствует *обыкновенное перемножение* характеристических функций.

Нетрудно сообразить, что если  $y = \sum_i x_i$  и каждое слагаемое независимо от суммы предшествующих, то

$$\varphi_y(u) = \prod_i \varphi_{x_i}(u). \quad (9.6)$$

Разумеется, это и подавно верно для  $x_i$ , независимых в совокупности.

Другой важный вопрос заключается в том, как по характеристической функции  $\varphi_x(u)$  найти функцию распределения  $W(x)$ .

Если распределение дифференцируемо, т. е.  $dW(x) = w(x)dx$ , то формула (9.1) переходит в обычный интеграл Фурье:

$$\varphi_x(u) = \int_{-\infty}^{+\infty} e^{iux} w(x) dx, \quad (9.7)$$

обращение которого тотчас же дает

$$w(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iux} \varphi_x(u) du, \quad (9.8)$$

а значит,

$$W(x) - W(0) = \int_0^x w(x) dx = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{1 - e^{-iux}}{iu} \varphi_x(u) du. \quad (9.9)$$

Пусть теперь  $w(x)$  — дискретное распределение. Пользуясь дельта-функцией, можно применять те же формулы обращения. Имеем

$$w(x) = \sum_k p_k \delta(x - a_k)$$

и, следовательно,

$$\varphi_x(u) = \int_{-\infty}^{+\infty} e^{iux} \sum_k p_k \delta(x - a_k) dx = \sum_k e^{iua_k} p_k.$$

Формула обращения тотчас же дает

$$\begin{aligned} w(x) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iux} \sum_k p_k e^{iua_k} du = \\ &= \sum_k p_k \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{iu(a_k - x)} du = \sum_k p_k \delta(x - a_k), \end{aligned}$$

т. е. и в этом случае возвращает нас к правильному выражению для  $w(x)$ .

Именно эта однозначная связь между функцией распределения и характеристической функцией широко используется как в общей теории, так и при решении конкретных задач. В одних случаях оказывается проще находить непосредственно функцию распределения, в других — характеристическую функцию. Укажем в этой связи на теорему Леви о том, что если последовательность функций распределения  $W_k(x)$  сходится при  $k \rightarrow \infty$  к функции распределения  $W(x)$  [в точках непрерывности  $W(x)$ ], то и последовательность соответствующих характеристических функций  $\varphi_k(u)$  сходится к характеристической функции  $\varphi(u)$ , отвечающей  $W(x)$ , и обратно.

Если взять  $\ln \varphi_x(u)$  и разложить его в степенной ряд по  $u$  [беря ту ветвь логарифма, которая вместе с  $u$  обращается в нуль, и разлагая в соответствии с (9.2) тоже до члена

порядка  $u^k$ ], то получим

$$\ln \varphi_x(u) = \sum_{m=1}^{\infty} \frac{\lambda_m}{m!} (iu)^m + o(u^k).$$

Величины  $\lambda_m$  называются *кумулянтами* (или *семиинвариантами*) распределения. Кумулянт  $m$ -го порядка есть целая рациональная функция первых  $m$  моментов  $\alpha_m$  (или центральных моментов  $\gamma_m$ ). В частности,

$$\lambda_1 = \alpha_1, \quad \lambda_2 = \alpha_2 - \alpha_1^2 = \gamma_2, \quad \lambda_3 = \alpha_3 - 3\alpha_2\alpha_1 + 2\alpha_1^3 = \gamma_3,$$

$$\lambda_4 = \alpha_4 - 3\alpha_2^2 - 4\alpha_3\alpha_1 + 12\alpha_2\alpha_1^2 - 6\alpha_1^4 = \gamma_4 - 3\gamma_2^2.$$

Следует обратить внимание на аддитивность кумулянтов при композиции распределений независимых величин. При композиции  $\varphi_{x_k}$  перемножаются,  $\ln \varphi_{x_k}$  складываются, а значит, складываются и кумулянты:

$$\lambda_m = \sum_{k=1}^n \lambda_m^{(k)}, \quad x = \sum_{k=1}^n x_k.$$

При помощи характеристических функций чрезвычайно просто устанавливается сохранение или несохранение вида некоторых распределений при композиции. Если, например, независимые случайные величины  $n_k$  распределены по Пуассону, то

$$\varphi_{n_k}(u) = \exp \{ \bar{n}_k (e^{iu} - 1) \}$$

и, следовательно, для  $n = \sum_k n_k$  имеем

$$\varphi_n(u) = \exp \{ \bar{n} (e^{iu} - 1) \}, \quad \bar{n} = \sum_k \bar{n}_k,$$

т. е. опять распределение Пуассона. Если сумма независимых одинаково распределенных случайных величин обладает тем же распределением, что и слагаемые, то такое распределение называется *устойчивым*. Таким образом, мы показали, что распределение Пуассона является устойчивым. Легко убедиться, что нормальное распределение тоже устойчиво.

Пусть независимые слагаемые распределены нормально со средними значениями  $\bar{x}_k = a_k$  и дисперсиями  $D[x_k] = \sigma_k^2$ , т. е.

$$W_k(x) = \Phi\left(\frac{x - a_k}{\sigma_k}\right), \quad \text{где} \quad \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-x^2/2} dx.$$

Характеристическая функция для  $x_k$  есть

$$\varphi_k(u) = e^{iua_k - \sigma_k^2 u^2/2}.$$



Это значит, что для  $\sum_{k=1}^n x_k$  характеристическая функция будет

$$\varphi(u) = \prod_{k=1}^n \varphi_k(u) = e^{iuA_n - B_n^2 u^2/2}, \quad \text{где } A_n = \sum_{k=1}^n a_k, \quad B_n^2 = \sum_{k=1}^n \sigma_k^2,$$

и, таким образом,

$$\begin{aligned} W_1(x) * W_2(x) * \dots * W_n(x) = \\ = \Phi\left(\frac{x-a_1}{\sigma_1}\right) * \dots * \Phi\left(\frac{x-a_n}{\sigma_n}\right) = \Phi\left(\frac{x-A_n}{B_n}\right). \end{aligned}$$

Если все  $W_k(x)$  одинаковы, т. е.  $a_k = a$ ,  $\sigma_k = \sigma$ , то  $A_n = na$ ,  $B_n = \sqrt{n} \sigma$ . Обозначая композицию одинаковых нормальных функций распределения  $W(x)$  через  $[W(x)]^{n*}$ , имеем

$$[W(x)]^{n*} = \left[ \Phi\left(\frac{x-a}{\sigma}\right) \right]^{n*} = \Phi\left(\frac{x-na}{\sqrt{n}\sigma}\right).$$

В частности, при  $a=0$  и  $\sigma=1/\sqrt{n}$

$$[\Phi(x\sqrt{n})]^{n*} = \Phi(x).$$

Это значит, что если  $x_k$  имеют одну и ту же функцию распределения  $\Phi(x)$ , то и сумма

$$S_n = \frac{1}{\sqrt{n}} \sum_{k=1}^n x_k$$

тоже имеет функцию распределения  $\Phi(x)$ .

Иногда оказывается удобным пользоваться не характеристической функцией — средним значением от  $e^{iux}$ , — а средними значениями от величин  $e^{\theta x}$ :

$$F_x(\theta) = \overline{e^{\theta x}} = \int e^{\theta x} dW(x),$$

или от  $u^x$ :

$$F_x(u) = \overline{u^x} = \int u^x dW(x),$$

разумеется, при ограничениях, обеспечивающих существование этих средних. Такого рода функции называются *моментопроизводящими* или просто *производящими* функциями распределения  $W(x)$ . Если возможно  $t$ -кратное дифференцирование  $F_x(\theta)$  по  $\theta$ , то, полагая затем  $\theta = 0$ , получаем  $\alpha_m = x^m$ . Аналогичным образом можно получать  $\alpha_m$  из  $F_x(u)$ , применяя  $t$  раз операцию и  $\frac{\partial}{\partial u}$  и полагая потом  $u = 1$ .

### § 10. Функция распределения импульсного пуассоновского процесса

Нахождение характеристической функции, будучи равносильно нахождению функции распределения, оказывается более простой задачей и для случайной величины (8.9):

$$\xi_n(t) = \sum_{v=1}^n x_v, \quad x_v = F(t - t_v, a_v),$$

так как слагаемые  $x_v$  независимы. Поэтому характеристическая функция (8.9) [т. е. условная функция — при условии, что в интервале  $(-T/2, T/2)$  было  $n$  импульсов] равна

$$\varphi_\xi(u|n) = \prod_{v=1}^n \varphi_{x_v}(u).$$

В свою очередь характеристическая функция для  $x_v$  (очевидно, не зависящая от номера импульса  $v$ ) есть

$$\varphi_{x_v}(u) = e^{iux_v} = \int w_a(a) da \int_{-T/2}^{T/2} e^{iuF(t-\tau, a)} w_t(\tau) d\tau.$$

Мы не пишем здесь пределов в интегралах по  $a$ , подразумевая, что интегрирование по  $a$  распространяется на всю область возможных значений (конечную или бесконечную) компонент этой  $m$ -мерной случайной величины. Что касается распределения моментов возникновения импульсов  $w_t(\tau) d\tau$ , то теперь, конечно, нельзя воспользоваться формулой (8.4), так как речь идет об импульсах, появляющихся не где угодно на оси  $t$ , а обязательно (с достоверностью) внутри интервала  $(-T/2, T/2)$ . Это значит, что равномерное распределение  $\tau$  должно быть нормировано к единице на этом интервале:

$$w_t(\tau) d\tau = \frac{n_1 d\tau}{\int_{-T/2}^{T/2} n_1 d\tau} = \frac{d\tau}{T}. \quad (10.1)$$

Таким образом,

$$\varphi_\xi(u|n) = \left\{ \int w_a(a) da \int_{-T/2}^{T/2} e^{iuF(t-\tau, a)} \frac{d\tau}{T} \right\}^n. \quad (10.2)$$

Обращая по Фурье формулу (8.7) [или полагая в (8.8)  $f(x) = e^{iux}$ ], мы получаем соотношение между безусловной и условной характеристическими функциями:

$$\varphi_\xi(u) = \sum_{n=0}^{\infty} P(n) \varphi_\xi(u|n). \quad (10.3)$$

Подстановка в (10.3) выражений (8.5) для  $P(n)$  и (10.2) для  $\varphi_{\xi}(u|n)$  дает

$$\begin{aligned}\varphi_{\xi}(u) &= \sum_{n=0}^{\infty} \frac{e^{-n_1 T}}{n!} \left\{ n_1 \int w_a(a) da \int_{-T/2}^{T/2} e^{iuF(t-\tau, a)} d\tau \right\}^n = \\ &= \exp \left\{ n_1 \int w_a(a) da \int_{-T/2}^{T/2} e^{iuF(t-\tau, a)} d\tau - n_1 T \right\},\end{aligned}$$

что можно записать также в виде

$$\begin{aligned}\varphi_{\xi}(u) &= \exp \left\{ n_1 \int w_a(a) da \int_{-T/2}^{T/2} [e^{iuF(t-\tau, a)} - 1] d\tau \right\} = \\ &= \exp \left\{ n_1 \int w_a(a) da \int_{t-T/2}^{t+T/2} [e^{iuF(\theta, a)} - 1] d\theta \right\}.\end{aligned}$$

В последнем выражении введена новая переменная интегрирования  $\theta = t - \tau$ .

Заметим теперь, что скобка под интегралом по  $\theta$  отлична от нуля только при тех  $\theta$ , для которых  $F(\theta, a) \neq 0$ , т. е. в пределах импульса. Поэтому для всех  $t$ , отступающих от краев интервала  $(-T/2, T/2)$  не менее чем на длительность импульса  $\theta$  (пренебрежение краевыми эффектами), можно раздвинуть пределы интегрирования по  $\theta$  в  $\pm \infty$ , или, другими словами, устремить  $T$  в бесконечность. Далее мы будем делать такой переход, не огораживая его каждый раз заново, но следует помнить, что он возможен лишь при достаточно быстром убывании  $F(\theta, a)$  при  $|\theta| \rightarrow \infty$ . В результате

$$\varphi_{\xi}(u) = \exp \left\{ n_1 \int w_a(a) da \int_{-\infty}^{+\infty} [e^{iuF(\theta, a)} - 1] d\theta \right\}. \quad (10.4)$$

Итак, характеристическая функция  $\xi(t)$  получена. В показателе экспоненты в (10.4) стоит логарифм этой характеристической функции. В его разложении по степеням  $iu$  коэффициентами при  $(iu)^m/m!$  являются, как мы знаем, кумулянты распределения  $\xi(t)$ . Разлагая этот показатель в ряд по степеням  $iu$ , получаем, что  $m$ -й кумулянт есть

$$\lambda_m = n_1 \int w_a(a) da \int_{-\infty}^{+\infty} F^m(\theta, a) d\theta. \quad (10.5)$$

В частности, для первого и второго кумулянтов, т. е. для среднего значения и дисперсии  $\xi(t)$ , имеем

$$\begin{aligned}\lambda_1 &= \overline{\xi(t)} = n_1 \int w_a(a) da \int_{-\infty}^{+\infty} F(\theta, a) d\theta, \\ \lambda_2 &= D[\xi(t)] = n_1 \int w_a(a) da \int_{-\infty}^{+\infty} F^2(\theta, a) d\theta.\end{aligned}\quad (10.6)$$

В случае, когда  $a_v$  — одномерная случайная величина, представляющая собой просто «амплитуду» импульса [процесс вида (8.1)], формулы (10.6) принимают вид

$$\overline{\xi(t)} = n_1 \bar{a} \int_{-\infty}^{+\infty} F(\theta) d\theta, \quad D[\xi(t)] = n_1 \bar{a}^2 \int_{-\infty}^{+\infty} F^2(\theta) d\theta. \quad (10.7)$$

Заметим, что  $\bar{\xi} = 0$  либо при  $\bar{a} = 0$ , либо при нулевой площади импульса  $\left(\int_{-\infty}^{+\infty} F(\theta) d\theta = 0\right)$ . Наконец, если «амплитуда» фиксирована, скажем,  $a_v = 1$  [другими словами, плотность вероятности  $a_v$  есть  $w_a(a_v) = \delta(a_v - 1)$ ], то мы приходим к формулам, первоначально полученным Кембеллом [1]:

$$\overline{\xi(t)} = n_1 \int_{-\infty}^{+\infty} F(\theta) d\theta, \quad D[\xi(t)] = n_1 \int_{-\infty}^{+\infty} F^2(\theta) d\theta. \quad (10.8)$$

Как характеристическая функция (10.4), так и кумулянты (10.5) содержат единственную характеристику «густоты» импульсов, а именно  $n_1$  — среднее число импульсов в единицу времени. Кроме того, все указанные величины не зависят от времени  $t$ . В этом проявляется *стационарность* рассматриваемого процесса, о чем еще будет идти речь в дальнейшем.

Разумеется, полученные формулы охватывают элементарную теорию дробового эффекта, которая была развита ранее (§ 5). Пусть отдельные импульсы одинаковы ( $a = e = \text{const}$ ), и пусть форма импульсов — прямоугольник продолжительности  $\theta$  и высоты  $1/\theta$  (рис. 3). Таким образом,

$$\int_{-\infty}^{+\infty} F(\theta) d\theta = 1, \quad \int_{-\infty}^{+\infty} F^2(\theta) d\theta = \frac{1}{\theta}.$$

Согласно (10.7) для импульсного тока  $I(t) = e \sum_v F(t - t_v)$  получаем

$$\bar{I} = n_1 e, \quad D[I] = \frac{n_1 e^2}{\theta} = \frac{e \bar{I}}{\theta}. \quad (10.9)$$

На первый взгляд выражение для  $D[I]$  находится в противоречии с полученным ранее [формула (5.2)], так как теперь в знаменателе стоит длительность импульса  $\vartheta$ , т. е. величина порядка времени пролета электрона, тогда как в (5.2) в знаменателе находилось время наблюдения  $T \gg \vartheta$ . В действительности никакого противоречия нет, так как формула (10.9) дает дисперсию *мгновенного* тока  $I(t)$ , в то время как (5.2) — это дисперсия тока, уже усредненного по промежутку времени  $T$ , т. е. дисперсия величины

$$I_T = \frac{1}{T} \int_0^T I(t) dt. \quad (10.10)$$

Естественно, что флуктуации  $I_T$  сглажены по сравнению с флуктуациями  $I(t)$ , и тем сильнее, чем больше  $T$ . Нетрудно убедиться в том, что для  $D[I_T]$  получается прежнее выражение (5.2) (см. задачу 6).

Полагая  $a_v = e = \text{const}$ , мы получали модель дробового эффекта. Если же значения случайного параметра  $a_v$  кратны заряду электрона, то процесс

$$\xi(t) = \sum_v e m_v \delta(t - t_v)$$

может при подходящем распределении вероятностей для  $m_v$  описать ток вторичной электронной эмиссии и его флуктуации:  $v$ -й электрон первичного тока  $I(t) = e \sum_v \delta(t - t_v)$  порождает слу-

чайное число  $m_v = 0, 1, 2, \dots$  вторичных электронов. Вместо формул (10.9) мы получим при этом для моментов  $\xi(t) \equiv I_{\text{вт.эм}}$

$$\bar{I}_{\text{вт. эм}} = n_1 \bar{m} e,$$

$$(\overline{\Delta I_{\text{вт. эм}}})^2 = \frac{n_1 \bar{m}^2 e^2}{\vartheta} = \frac{e \bar{I}_{\text{вт. эм}}}{\vartheta} \frac{\bar{m}^2}{\bar{m}}.$$

Зная характеристическую функцию (10.4), мы в принципе знаем и ее трансформанту Фурье, т. е. плотность вероятности  $\omega(x)$  процесса  $\xi(t)$ . Однако вычисление соответствующего интеграла Фурье в замкнутом виде осуществимо лишь в немногих простейших случаях. Если, например, процесс состоит из одинаковых прямоугольных импульсов фиксированной амплитуды  $a$  и длительности  $\vartheta$ :

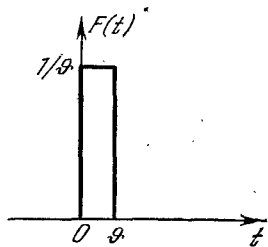


Рис. 3.

$$F(\theta, a) = a F(\theta), \quad F(\theta) = \begin{cases} 1 & \text{в интервале } (0, \vartheta), \\ 0 & \text{вне этого интервала,} \end{cases} \quad (10.11)$$

то (10.4) дает

$$\varphi_{\xi}(u) = \exp \left\{ n_1 \int_0^{\theta} (e^{iua} - 1) d\theta \right\} = \exp \{ n_1 \theta (e^{iua} - 1) \}.$$

Но это характеристическая функция распределения Пуассона с параметром  $\bar{n} = n_1 \theta$  и дискретными возможными значениями  $na$ , где  $n = 0, 1, 2, \dots$ :

$$P \{ \xi(t) = na \} = \frac{(n_1 \theta)^n}{n!} e^{-n_1 \theta}. \quad (10.12)$$

В общем же случае решенная задача как раз такова, когда целесообразно искать не функцию распределения, а характеристическую функцию, что мы и сделали выше.

Однако при определенных условиях можно установить, что (10.4) принимает некоторый универсальный вид. Выясним, каков этот вид и в чем состоят упомянутые условия.

Обращая характеристическую функцию

$$\varphi_{\xi}(u) = \exp \left\{ \sum_1^{\infty} \lambda_m \frac{(iu)^m}{m!} \right\}$$

по Фурье, находим плотность вероятности

$$\begin{aligned} w(\xi) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iu\xi} \varphi_{\xi}(u) du = \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp \left\{ -iu\xi + \sum_{m=1}^{\infty} \lambda_m \frac{(iu)^m}{m!} \right\} du. \end{aligned}$$

Если ввести обозначения

$$\lambda_1 = \bar{\xi}, \quad \lambda_2 = \sigma_{\xi}^2 = D[\xi(t)], \quad x = \frac{\xi - \lambda_1}{\sqrt{\lambda_2}},$$

$$\Phi^{(n)}(x) = \frac{1}{\sqrt{2\pi}} \frac{d^n e^{-x^2/2}}{dx^n}$$

и, начиная с  $m = 3$ , разложить экспоненту под интегралом по степеням  $iu$ , то почленное интегрирование приводит к представлению  $w(\xi)$  в виде ряда по производным  $\Phi^{(n)}(x)$ . Собрав в этом ряде члены одного порядка относительно  $n_1$ , мы получаем так называемый ряд Эджворта<sup>1)</sup>:

$$w(\xi) = \frac{\Phi^{(0)}(x)}{\sigma_{\xi}} - \frac{\lambda_3 \Phi^{(3)}(x)}{3! \sigma_{\xi}^4} + \left[ \frac{\lambda_4 \Phi^{(4)}(x)}{4! \sigma_{\xi}^5} + \frac{\lambda_3^2 \Phi^{(6)}(x)}{72 \sigma_{\xi}^7} + \dots \right]. \quad (10.13)$$

<sup>1)</sup> См. [2], гл. 17, § 7.

Первый член здесь порядка  $n_1^{-1/2}$ , второй — порядка  $n_1^{-1}$ , третий —  $n_1^{-3/2}$  и т. д. Все члены пропорциональны  $e^{-x^2/2}$ , а первый член есть просто нормальное распределение для  $x$ . Очевидно, с ростом  $n_1$ , т. е. с увеличением «густоты» импульсов, первый член будет все более преобладать, т. е. распределение  $w(\xi)$  будет приближаться к нормальному.

Сделаем грубую оценку того значения  $n_1$ , при котором уже можно ограничиться первым членом ряда (10.13). Два первых члена ряда в раскрытой форме таковы:

$$w(\xi) = \frac{e^{-x^2/2}}{\sqrt{2\pi}\sigma_\xi} \left\{ 1 - \frac{\lambda_3}{6\sigma_\xi^3} x(3-x^2) \right\}.$$

Наличие экспоненциального множителя позволяет не рассматривать большие значения  $x$ , при которых  $w(\xi)$  уже очень мало. Если ограничиться  $x \leq 2$ , то  $|x(3-x^2)| \leq 2$ . Значит, условие малости второго члена будет

$$\frac{|\lambda_3|}{3\sigma_\xi^3} \leq \frac{n_1 |\overline{a^3}| \int_{-\infty}^{+\infty} |F(\theta)|^3 d\theta}{3 \left[ n_1 \overline{a^2} \int_{-\infty}^{+\infty} F^2(\theta) d\theta \right]^{3/2}} \ll 1$$

или, если отбросить коэффициент порядка единицы,

$$\frac{\int_{-\infty}^{+\infty} |F(\theta)|^3 d\theta}{n_1^{1/2} \left[ \int_{-\infty}^{+\infty} F^2(\theta) d\theta \right]} \ll 1.$$

Если  $\int_{-\infty}^{+\infty} F(\theta) d\theta = 1$  и длительность импульса порядка  $\theta$ , то интеграл в числителе будет порядка  $1/\theta^2$ , а в знаменателе — порядка  $1/\theta$ , так что написанное условие сведется к следующему:

$$n_1 \theta \gg 1. \quad (10.14)$$

Это означает, что число импульсов, возникающих за длительность одного импульса, должно быть достаточно велико. Можно выразить то же самое и несколько иначе. В момент времени  $t$  величина  $\xi(t)$  складывается из тех импульсов, которые возникли в интервале времени от  $t-\theta$  до  $t$ , так как более ранние импульсы к моменту  $t$  уже успели «затухнуть». Следовательно,  $n_1 \theta$  — это в среднем число слагаемых  $x_v$ , из которых состоит  $\xi(t)$ .

в каждый данный момент. Условие (10.14) говорит, таким образом, о том, что распределение  $\xi(t)$  тем ближе к нормальному, чем больше импульсов налагается в каждый момент времени. Если толчки редкие, то распределение будет существенно зависеть от формы индивидуального толчка, нормального распределения не будет. Прийти к нему можно либо учащая толчки, либо увеличивая длительность импульса. В обоих случаях число толчков за время одного импульса будет расти.

Этот результат представляет собой весьма частный случай так называемой *центральной предельной теоремы* теории вероятностей. По сути дела, здесь было воспроизведено доказательство этой теоремы, но применительно к очень специальным условиям рассмотренной задачи. Мы увидим в дальнейшем (§ 13), что и найденный результат, и условие (10.14) непосредственно вытекают из гораздо более общих положений.

В работе Джильберта и Поллака [3] распределение импульсного процесса (8.1) было найдено для некоторых форм импульса  $F(t)$  в явном виде и в общем случае, т. е. без предельного перехода к «густому» шуму ( $n_1\theta \rightarrow \infty$ ). Во-первых, авторы показали, что процесс (8.1) и даже более общий процесс

$$\xi(t) = \sum_{\nu} F(t - t_{\nu}, a_{\nu}) \quad (10.15)$$

всегда можно свести к шуму вида

$$\xi(t) = \sum_{\nu} F_{\text{ЭКВ}}(t - t_{\nu}) \quad (10.16)$$

с такой формой импульса  $F_{\text{ЭКВ}}(t)$ , что распределение

$$W(\xi) = \int_{-\infty}^{\xi} w(x) dx$$

будет для (10.15) и (10.16) одним и тем же. Во-вторых, для  $W(\xi)$  они вывели (двумя различными способами) интегральное уравнение

$$\int_{-\infty}^{\xi} x dW(x) = n_1 \int_{-\infty}^{+\infty} W[\xi - F(\theta)] F(\theta) d\theta,$$

которое и сумели решить для некоторых форм импульса  $F(t)$ .

Конечно, для прямоугольных импульсов высоты  $a$  и длительности  $\theta$  этому уравнению удовлетворяет распределение Пуассона (10.12), которое, как мы помним, переходит с ростом  $n_1\theta$  в распределение Лапласа, т. е. при сглаживании ступенек высоты  $a$  — в нормальное распределение.



Если форма импульса задана следующим несколько искусственным образом (рис. 4; через  $\gamma$  на нем обозначена постоянная Эйлера):

$$F(t) = 0 \quad \text{при } t \leq 0,$$

$$F(t) = \int_0^\infty \frac{e^{-y}}{y} dy \quad \text{при } t > 0,$$

то распределение  $\xi$  получается экспоненциально-степенным («непрерывным пуассоновским»):

$$w(x) = \frac{x^{\alpha-1} e^{-x}}{\Gamma(\alpha)} \quad (0 < x < \infty), \quad (10.17)$$

где  $\alpha = \bar{\xi}$ . С ростом  $\alpha$  оно также переходит в нормальное.

К сожалению, для такой важной формы импульса, как затухающее колебание (импульсная функция резонансного контура), уравнение для  $W$  удастся решить только численными методами, но и такую возможность следует рассматривать как определенное достижение.

Отметим, что (10.17) представляет собой частный случай так называемого *гамма-распределения*:

$$w(x) = \frac{\kappa^\alpha x^{\alpha-1}}{\Gamma(\alpha)} e^{-\kappa x} \quad (0 < x < \infty), \quad (10.18)$$

характеристическая функция которого

$$\varphi(u) = \left( \frac{\kappa}{\kappa - iu} \right)^\alpha, \quad (10.19)$$

а среднее значение и дисперсия следующим образом выражаются через положительные параметры  $\alpha$  и  $\kappa$ :

$$\bar{x} = \frac{\alpha}{\kappa}, \quad D[x] \equiv \sigma_x^2 = \frac{\alpha}{\kappa^2}.$$

Очевидно, (10.17) следует из (10.18), если параметр  $\kappa$  (называемый масштабным) положить равным единице. При целочисленных значениях  $\alpha$  ( $\alpha = n$ ) имеем из (10.18)

$$w(x) = \frac{\kappa^n}{(n-1)!} x^{n-1} e^{-\kappa x}.$$

Обычно это распределение записывают через параметр  $\mu = 1/\bar{x}$  (так что  $\kappa = n\mu$  и  $\sigma_x^2 = 1/n\mu^2$ ):

$$w(x) = \frac{(n\mu)^n}{(n-1)!} x^{n-1} e^{-n\mu x}. \quad (10.20)$$

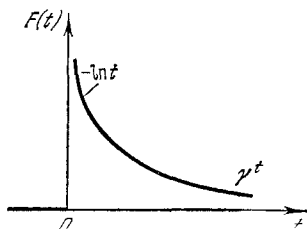


Рис. 4.

## § 11. Корреляционная функция

Рассматривая  $\xi(t)$  как случайную величину, мы нашли в предыдущем параграфе характеристическую функцию (10.4) и тем самым кумулянты  $\xi(t)$ , в частности среднее значение и дисперсию (10.6). Но для вычисления смешанных моментов двух случайных величин  $\xi_1 = \xi(t_1)$  и  $\xi_2 = \xi(t_2)$ , т. е. значений случайного процесса  $\xi(t)$  в два *различных* момента времени, надо знать совместную (двумерную) функцию распределения  $\xi_1$  и  $\xi_2$  или найти соответствующую двумерную характеристическую функцию. Однако мы не будем решать эту задачу, а ограничимся вычислением только смешанного момента  $\xi_1 \xi_2 - \bar{\xi}_1 \cdot \bar{\xi}_2$ , который является одной из важнейших характеристик случайных процессов, поскольку он определяет *корреляцию* между величинами  $\xi_1$  и  $\xi_2$ . В общем случае этот момент зависит от  $t_1$  и  $t_2$ , или, если положить  $t_1 = t$ ,  $t_2 = t + \tau$ , — от  $t$  и  $\tau$ ; он называется *корреляционной функцией* случайного процесса  $\xi(t)$ :

$$\psi_{\xi}(t, \tau) = \overline{\xi(t) \xi(t + \tau)} - \bar{\xi}(t) \cdot \bar{\xi}(t + \tau). \quad (11.1)$$

Так как кумулянт  $\lambda_1 = \bar{\xi}(t)$  оказался не зависящим от  $t$  (как и все  $\lambda_m$ ), второй член в правой части (11.1) равен  $\lambda_1^2$ . Следовательно, нам надо вычислить только смешанный момент

$$B_{\xi}(t, \tau) = \overline{\xi(t) \xi(t + \tau)}. \quad (11.2)$$

Схема расчета та же, что и в одномерном случае.

Запишем двумерную плотность вероятности для случайных величин

$$\begin{aligned} \xi &\equiv \xi(t) = \sum_{\nu} F(t - t_{\nu}, a_{\nu}), \\ \xi_{\tau} &\equiv \xi(t + \tau) = \sum_{\mu} F(t + \tau - t_{\mu}, a_{\mu}) \end{aligned} \quad (11.3)$$

через условные плотности вероятности [при условии, что интервал времени  $(-T/2, T/2)$  содержит  $n$  импульсов]:

$$w(x, y) = \sum_{n=0}^{\infty} P(n) v(x, y | n). \quad (11.4)$$

Отсюда следует, что

$$\overline{\xi \xi_{\tau}} = \sum_{n=0}^{\infty} P(n) \overline{\xi \xi_{\tau}}^{n-},$$

где

$$\overline{\xi \xi_{\tau}}^{n-} = \int_{-\infty}^{+\infty} x y v(x, y | n) dx dy$$

— *условный* момент, т. е. среднее произведение величин (11.3), но с суммированием, распространенным только на те  $n$  импульсов, которые содержатся в  $(-T/2, T/2)$ :

$$\overline{\xi\xi_\tau} = \sum_{\nu, \mu=1}^n \langle F(t - t_\nu, a_\nu) F(t + \tau - t_\mu, a_\mu) \rangle.$$

Усреднение производится здесь по всем случайным параметрам  $t_\nu, t_\mu, a_\nu, a_\mu$  ( $\nu, \mu = 1, 2, \dots, n$ ). Так как величины с  $\nu \neq \mu$  по предположению независимы, двойную сумму можно разбить на две части:

$$\begin{aligned} \overline{\xi\xi_\tau} &= \sum_{\nu=1}^n \langle F(t - t_\nu, a_\nu) F(t + \tau - t_\nu, a_\nu) \rangle + \\ &\quad + \sum_{\nu \neq \mu} \langle F(t - t_\nu, a_\nu) \rangle \langle F(t + \tau - t_\mu, a_\mu) \rangle. \end{aligned}$$

Но средние значения не зависят в силу наших исходных допущений от номеров импульсов  $\nu$  и  $\mu$ . Поэтому первая сумма содержит  $n$  одинаковых слагаемых, а вторая —  $(n^2 - n)$ :

$$\begin{aligned} \overline{\xi\xi_\tau} &= n \langle F(t - \chi, a) F(t + \tau - \chi, a) \rangle + \\ &\quad + (n^2 - n) \langle F(t - \chi, a) \rangle \langle F(t + \tau - \chi', a') \rangle, \end{aligned}$$

где

$$\begin{aligned} \langle F(t - \chi, a) F(t + \tau - \chi, a) \rangle &= \\ &= \int w_a(a) da \int_{-T/2}^{T/2} F(t - \chi, a) F(t + \tau - \chi, a) \frac{d\chi}{T}, \\ \langle F(t - \chi, a) \rangle &= \int w_a(a) da \int_{-T/2}^{T/2} F(t - \chi, a) \frac{d\chi}{T}, \\ \langle F(t + \tau - \chi', a) \rangle &= \int w_a(a) da \int_{-T/2}^{T/2} F(t - \tau - \chi', a) \frac{d\chi'}{T}. \end{aligned}$$

Сделав в двух первых выражениях замену переменной интегрирования  $t - \chi = \theta$ , а в третьем — замену  $t + \tau - \chi' = \theta$  и раздвигая затем пределы интегралов по  $\theta$  в  $\pm \infty$ , получаем

$$\begin{aligned} \overline{\xi\xi_\tau} &= \frac{T}{n} \int w_a(a) da \int_{-\infty}^{+\infty} F(\theta, a) F(\theta + \tau, a) d\theta + \\ &\quad + \frac{n^2 - n}{T} \left[ \int w_a(a) da \int_{-\infty}^{+\infty} F(\theta, a) d\theta \right]^2. \end{aligned}$$

Это условное среднее остается усреднить теперь по распределению Пуассона, для которого, как мы знаем,  $\bar{n} = n_1 T$  и  $D[n] = \bar{n}^2 - \bar{n}^2 = \bar{n}$  (т. е.  $\bar{n}^2 - \bar{n} = \bar{n}^2 = n_1^2 T^2$ ), и мы получаем, согласно (11.4), момент второго порядка

$$B_{\xi}(\tau) = \overline{\xi\xi_{\tau}} = n_1 \int w_a(a) da \int_{-\infty}^{+\infty} F(\theta, a) F(\theta + \tau, a) d\theta + \lambda_1^2,$$

где мы воспользовались первой формулой (10.6). Следовательно, корреляционная функция (11.1) есть

$$\psi_{\xi}(\tau) = n_1 \int w_a(a) da \int_{-\infty}^{+\infty} F(\theta, a) F(\theta + \tau, a) d\theta. \quad (11.5)$$

Итак, корреляционная функция (и момент  $B_{\xi}$ ) не зависит от времени  $t$ , а только от *разности*  $\tau$  обоих моментов времени. Это снова результат стационарности рассматриваемого случайного процесса. При  $\tau = 0$  формула (11.5) переходит в формулу для дисперсии  $\xi(t)$  [см. вторую формулу (10.6)]:

$$\psi_{\xi}(0) = D[\xi(t)] = \lambda_2,$$

как это и должно быть в соответствии с определением (11.1) корреляционной функции.

Нетрудно видеть, что  $\psi_{\xi}(\tau)$  — четная функция  $\tau$ . При  $|\tau| \rightarrow \infty$ , если  $F(\theta)$  достаточно быстро убывает при  $|\theta| \rightarrow \infty$  (быстрее, чем  $1/\sqrt{|\theta|}$ ),  $\psi_{\xi}(\tau)$  стремится к нулю. Ясно, что с ростом  $|\tau|$  сомножители подынтегрального выражения в (11.5) раздвигаются и при указанном условии произведение стремится к нулю быстрее, чем  $1/|\theta|$ .

В частном случае, когда  $a$  — одномерная случайная величина, а именно «амплитуда» импульса [ $F(\theta, a) = aF(\theta)$ ], получаем

$$\psi_{\xi}(\tau) = n_1 \overline{a^2} \int_{-\infty}^{+\infty} F(\theta) F(\theta + \tau) d\theta. \quad (11.6)$$

Одной из компонент  $n$ -мерной случайной величины  $a_v$  может быть и случайная длительность  $v$ -го импульса  $\theta_v$ . Пусть  $\theta_v$  независимо от остальных компонент  $a_v$ , так что распределение  $w_a(a_v) da_v$  должно быть заменено на  $w_a(a_v) da_v \cdot w_{\theta}(\theta_v) d\theta_v$ . Если мы хотим явным образом учесть даваемое  $\theta_v$  случайное изменение длительности импульсов, то целесообразно представить детерминированную функцию  $F$ , описывающую форму импульса, в виде  $F\left(\frac{t - t_v}{\theta_v}, a_v\right)$ . Весь вывод характеристической

функции, кумулянтов и корреляционной функции остается при этом таким же, как и выше, но при усреднении по  $t_v$  — случайным моментам возникновения импульсов — удобно ввести новую переменную интегрирования  $\theta$  следующим образом:  $\theta = (t - t_v)/\vartheta_v$ . Тогда распределение  $dt_v/T$  в  $(-T/2, T/2)$  перейдет в  $\vartheta_v d\theta/T$  в пределах от  $(t - T/2)/\vartheta_v$  до  $(t + T/2)/\vartheta_v$ . После раздвигания пределов в  $\pm\infty$  параметр  $\vartheta_v$  сохранится только множителем в подынтегральной функции (а при вычислении корреляционной функции — еще и как масштаб для временного сдвига  $\tau$ , т. е. вместо  $\tau$  будет  $\tau/\vartheta_v$ ). В результате для кумулянтов  $\lambda_m$  мы получим вместо (10.5) формулу

$$\lambda_m = n_1 \bar{\vartheta} \int_{-\infty}^{+\infty} \langle F^m(\theta, \mathbf{a}) \rangle d\theta, \quad (11.7)$$

а для корреляционной функции вместо (11.5) формулу

$$\psi_{\xi}(\tau) = n_1 \int_0^{\infty} \vartheta \omega_{\vartheta}(\vartheta) d\vartheta \int_{-\infty}^{+\infty} \langle F(\theta, \mathbf{a}) F(\theta + \frac{\tau}{\vartheta}, \mathbf{a}) \rangle d\theta, \quad (11.8)$$

где угловыми скобками обозначено для краткости усреднение по распределению  $\mathbf{a}$ .

Рассмотрим следующий пример. Пусть  $\mathbf{a}_v$  — одномерная случайная «амплитуда», так что

$$\xi(t) = \sum_v a_v F\left(\frac{t - t_v}{\vartheta_v}\right). \quad (11.9)$$

Пусть, далее,  $F(\theta) = 1$  в интервале  $(0, 1)$  и равно нулю вне этого интервала, причем длительности  $\vartheta$  распределены по закону

$$\omega(\vartheta) d\vartheta = e^{-\vartheta/\bar{\vartheta}} \frac{d\vartheta}{\bar{\vartheta}}.$$

Тогда при  $\tau > 0$

$$\int_{-\infty}^{+\infty} F(\theta) F\left(\theta + \frac{\tau}{\vartheta}\right) d\theta = \begin{cases} 0 & \text{при } \vartheta < \tau, \\ 1 - \frac{\tau}{\vartheta} & \text{при } \vartheta > \tau, \end{cases}$$

а значит,

$$\psi(\tau) = n_1 \bar{a}^2 \int_{\tau}^{\infty} (\vartheta - \tau) e^{-\vartheta/\bar{\vartheta}} \frac{d\vartheta}{\bar{\vartheta}} = n_1 \bar{a}^2 \bar{\vartheta} e^{-\tau/\bar{\vartheta}}.$$

Функция корреляции четна по  $\tau$ , так что при произвольном знаке  $\tau$  результат можно записать в виде

$$\psi(\tau) = n_1 \bar{a}^2 \bar{\vartheta} e^{-|\tau|/\bar{\vartheta}}. \quad (11.10)$$

Рис. 5 дает представление о характере этого процесса в частном случае фиксированных  $a_v = 1$ . Было высказано предположение, что такого рода процесс способен послужить моделью для эффекта мерцания — относительно медленных флуктуаций тока эмиссии, обусловленных локальными изменениями эмиссионной способности поверхности катода [4]. Если причиной этих изменений является присутствие на поверхности примесных атомов, в частности сающихся на катод атомов остаточного газа,

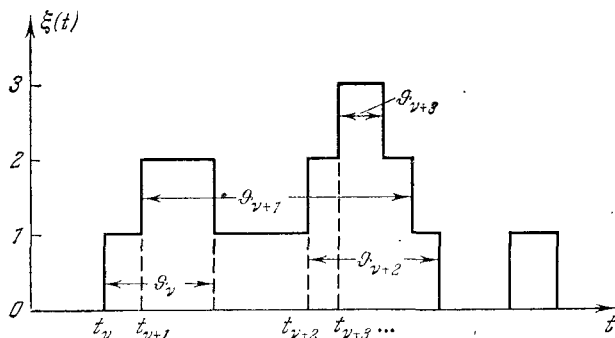


Рис. 5.

то  $\xi(t)$  — это случайное мгновенное число таких атомов, а  $\theta_v$  — времена их пребывания на поверхности. Однако подобная модель не согласуется с формой спектра эффекта мерцания, которая в широком диапазоне частот  $\omega$  большей частью близка к  $\omega^{-1}$  (см. § 55). Наряду с попытками усовершенствовать теорию введением набора средних времен пребывания  $\bar{\theta}$  с некоторыми весами или посредством других аналогичных допущений, была высказана мысль о том, что спектр требуемого вида может быть связан с процессом (8.1), если импульсы  $F(t)$  очень медленно спадают с ростом  $t^1$ , а именно, если при достаточно больших  $t$  функция  $F(t)$  убывает, как  $t^{\alpha/2-1}$  ( $0 < \alpha < 2$ ). Однако при  $\alpha \geq 1$  процесс уже не является стационарным. Мы вернемся к этому кругу вопросов в дальнейшем, при рассмотрении спектральных разложений случайных процессов.

Приведем в заключение еще один пример импульсного процесса, связанного с пуассоновским. Это — так называемый обобщенный телеграфный сигнал — чередующиеся скачки  $\xi(t)$  между значениями  $-a$  и  $+a$ , причем моменты скачков  $t_v$  распределены по Пуассону (рис. 6). Для простоты мы примем среднюю плотность скачков  $n_1$  постоянной. Добавление к  $\xi(t)$  постоянной составляющей  $\xi_0 > a$  дало бы односторонние импульсы, и тогда

<sup>1</sup>) См. [5], где приведен также обзор литературы.

рассматриваемый процесс мог бы служить очень упрощенной моделью флуктуаций проводимости в находящихся под током гранулированных сопротивлениях, представляющих собой конгломерат (чаще всего пленку) из хорошо проводящих контактирующих зерен. Поскольку постоянная составляющая  $\bar{\xi} = \xi_0$  не существенна для статистических свойств процесса, мы будем считать  $\bar{\xi} = 0$ . Распределение моментов  $t_v$  смены знака  $\xi$  по Пуассону означает, что длительности импульсов имеют экспоненциальное распределение

$$n_1 e^{-n_1 \tau} d\tau.$$

Казалось бы, что, рассматривая производную  $\dot{\xi}(t)$ , состоящую из импульсов  $\pm 2a\delta(t - t_v)$ , мы возвращаемся к уже разобранный задаче — процессу вида (8.1). Легко видеть, однако, что это не так, поскольку теперь является обязательным чередование знаков скачков, в то время как в (8.1) значения  $a_v$  были взаимно независимы.

Если в интервале  $(t, t + \tau)$  знак  $\xi(t)$  изменился четное число раз, то  $\xi(t + \tau)$  того же знака, что и  $\xi(t)$ , т. е.  $\xi(t + \tau)\xi(t) = a^2$ . В противном случае это произведение равно  $-a^2$ . Следовательно, при  $\tau > 0$

$$\psi(\tau) = a^2 P_{\text{чет}} - a^2 P_{\text{неч}},$$

где  $P_{\text{чет}}$  и  $P_{\text{неч}}$  — вероятности четного и нечетного числа скачков за время  $\tau$ . В соответствии с законом Пуассона

$$P(n) = \frac{\bar{n}^n}{n!} e^{-\bar{n}}, \quad \text{где } \bar{n} = n_1 \tau,$$

имеем

$$P_{\text{чет}} = P(0) + P(2) + \dots = e^{-\bar{n}} \left( 1 + \frac{\bar{n}^2}{2!} + \frac{\bar{n}^4}{4!} + \dots \right) = e^{-\bar{n}} \text{ch } \bar{n},$$

$$P_{\text{неч}} = P(1) + P(3) + \dots = e^{-\bar{n}} \left( \bar{n} + \frac{\bar{n}^3}{3!} + \frac{\bar{n}^5}{5!} + \dots \right) = e^{-\bar{n}} \text{sh } \bar{n}.$$

Следовательно, при  $\tau > 0$

$$\psi(\tau) = a^2 e^{-\bar{n}} (\text{ch } \bar{n} - \text{sh } \bar{n}) = a^2 e^{-2\bar{n}} = a^2 e^{-2n_1 \tau},$$

а при любом знаке  $\tau$

$$\psi(\tau) = a^2 e^{-2n_1 |\tau|}. \quad (11.11)$$

Отметим, что функция корреляции (11.11) совпадает по форме с (11.10), несмотря на совершенно различный вид реализации самих случайных процессов (рис. 5 и 6). Такая же функция

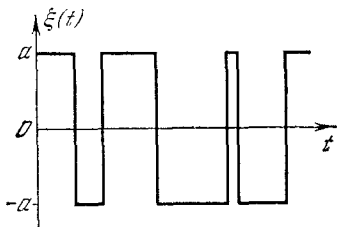


Рис. 6.

корреляции будет и у процесса (8.1), если он состоит из экспоненциальных импульсов

$$F(t) = \begin{cases} 0 & \text{при } t < 0, \\ e^{-\alpha t} & \text{при } t > 0. \end{cases}$$

Несложный расчет по формуле (11.6) показывает, что и в этом случае

$$\psi(\tau) = \frac{n_1 \bar{a}^2}{2\alpha} e^{-\alpha |\tau|}. \quad (11.12)$$

Представление о реализации такого процесса, который может являться откликом  $RC$ -цепочки на случайную последовательность дельта-импульсов, дает рис. 7.

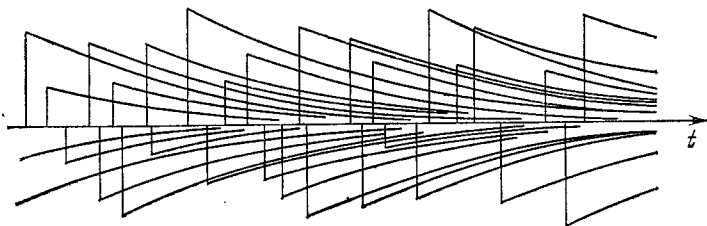


Рис. 7.

Функция корреляции вида  $Ae^{-\alpha|\tau|}$  часто употребляется в качестве аппроксимации в тех случаях, когда фактическая  $\psi(\tau)$  монотонно спадает, не меняя знака. Такая аппроксимация очень удобна (наряду с гауссовой кривой) для разного рода теоретических оценок и приближенных выводов.

## § 12. Некоторые обобщения задачи об импульсах

В основе предыдущих выводов лежало предположение о том, что плотность вероятности возникновения импульса в промежутке  $(t, t + dt)$  постоянна и, следовательно, не зависит ни от момента времени  $t$ , ни от наличия и положения других импульсов. Отказ от любого из этих допущений ведет к более общим случаям.

Если вероятность появления импульса в  $(t, t + dt)$  зависит от промежутков времени между  $t$  и моментами возникновения некоторого числа предшествующих импульсов, то это будет уже условная вероятность<sup>1)</sup>. Рассмотрим первоначально именно эту задачу, но ограничимся случаем, когда на вероятность появления  $v$ -го импульса влияет только  $(v-1)$ -й импульс, т. е. эта

<sup>1)</sup> Постановку и решение этой задачи дал Домб [6].



вероятность зависит от времени  $\vartheta$ , протекшего с момента появления  $(\nu - 1)$ -го импульса:

$$P \left\{ \begin{array}{l} \nu\text{-й импульс} \\ \text{в } (\vartheta, \vartheta + d\vartheta) \end{array} \middle| \begin{array}{l} \text{при условии, что} \\ \text{интервал } \vartheta \text{ пустой} \end{array} \right\} = \varphi(\vartheta) d\vartheta. \quad (12.1)$$

Естественно ожидать, что с ростом  $\vartheta$  влияние  $(\nu - 1)$ -го импульса ослабевает, т. е.  $\varphi(\vartheta) \rightarrow \text{const}$ .

Раньше у нас было  $\varphi(\vartheta) = n_1^0 = \text{const}$  при всяком  $\vartheta$  [т. е. распределение (12.1) было безусловным], и это давало для числа импульсов  $n$  в интервале  $(0, T)$  распределение Пуассона<sup>1)</sup>. Теперь распределение  $n$ , конечно, не будет пуассоновским, и соответствующее изменение моментов  $\bar{n}$  и  $\bar{n}^2$  будет интересовать нас в первую очередь.

Как связана с  $\varphi(\vartheta)$  вероятность  $W(\vartheta)$  того, что промежуток  $\vartheta$  после  $(\nu - 1)$ -го импульса пустой? Очевидно, вероятность  $W(\vartheta + d\vartheta)$  пустого промежутка  $\vartheta + d\vartheta$  равна произведению вероятности  $W(\vartheta)$  пустого  $\vartheta$  на условную вероятность того, что при пустом  $\vartheta$  интервал  $(\vartheta, \vartheta + d\vartheta)$  тоже пустой. «Не пустой» интервал  $d\vartheta$  означает, что в нем может возникнуть не только  $\nu$ -й импульс, но и  $(\nu + 1)$ -й,  $(\nu + 2)$ -й и т. д. Мы уже видели, однако, что вероятность появления  $m$  импульсов в бесконечно малом интервале  $d\vartheta$  является величиной порядка  $(d\vartheta)^m$ . Поэтому условная вероятность пустого  $d\vartheta$  с точностью до первого порядка малости по  $d\vartheta$  равна  $1 - \varphi(\vartheta)d\vartheta$ . Таким образом,

$$W(\vartheta + d\vartheta) = W(\vartheta) [1 - \varphi(\vartheta) d\vartheta].$$

Ограничиваясь и слева первым порядком относительно  $d\vartheta$  и выполнив интегрирование, находим

$$W(\vartheta) = \exp \left\{ - \int_0^{\vartheta} \varphi(\theta) d\theta \right\}. \quad (12.2)$$

Введем, далее, вероятность  $g(\vartheta)d\vartheta$  того, что между двумя последовательными импульсами протекло время от  $\vartheta$  до  $\vartheta + d\vartheta$ . Эта вероятность равна произведению вероятности  $W(\vartheta)$  пустого  $\vartheta$  на условную вероятность (12.1) появления импульса в  $(\vartheta, \vartheta + d\vartheta)$  при пустом  $\vartheta$ :

$$g(\vartheta) d\vartheta = W(\vartheta) \varphi(\vartheta) d\vartheta = \exp \left\{ - \int_0^{\vartheta} \varphi(\theta) d\theta \right\} \varphi(\vartheta) d\vartheta. \quad (12.3)$$

Обратимся теперь к непосредственно интересующей нас вероятности  $P_n(T)$  появления в интервале  $(0, T)$   $n$  импульсов. Со-

<sup>1)</sup> Верхний индекс 0 при  $n_1$  введен теперь специально, чтобы выделить случай некоррелированных импульсов.

бытие « $n$  импульсов в  $(0, T)$ » можно разбить на взаимно исключающие частные случаи, состоящие в том, что последний ( $n$ -й) импульс возник за время  $(\eta, \eta + d\eta)$  до конца интервала  $(0, T)$ . Обозначив вероятность такого частного случая через  $p_n(T, \eta)d\eta$ , имеем

$$P_n(T) = \int_0^T p_n(T, \eta) d\eta.$$

Но  $p_n(T, \eta)d\eta$  можно записать как произведение вероятности  $W(\eta)$  того, что интервал  $\eta$  пустой (т. е.  $n$ -й импульс последний), на вероятность того, что  $n$ -й импульс появился в конце интервала  $(0, T - \eta)$ , равную  $p_n(T - \eta, 0)d\eta$ :

$$p_n(T, \eta) = W(\eta) p_n(T - \eta, 0).$$

Следовательно,

$$P_n(T) = \int_0^T W(\eta) p_n(T - \eta, 0) d\eta. \quad (12.4)$$

Вероятности же  $p_n(t, 0) \equiv w_n(t)$  удовлетворяют следующему интегрально-разностному уравнению:

$$w_n(T) = \int_0^T w_{n-1}(T - \eta) g(\eta) d\eta. \quad (12.5)$$

Действительно, под интегралом мы имеем произведение вероятностей того, что  $(n - 1)$ -й импульс появился в конце интервала  $(0, T - \eta)$ , и того, что между  $(n - 1)$ -м и  $n$ -м импульсами прошло время  $\eta$ , т. е.  $n$ -й импульс оказался в конце интервала  $(0, T)$ . Разные значения  $\eta$  дают взаимно исключающие частные случаи события « $n$ -й импульс в конце  $(0, T)$ », полная вероятность которого и записана в (12.5).

Решая (12.5) при заданных  $g(\eta)$  и  $w_1(t)$ , мы получаем  $w_n(t) \equiv p_n(t, 0)$ , что позволяет при помощи (12.4) вычислить  $P_n(T)$ . Однако прямое осуществление этой программы довольно сложно.

В цитируемой работе [6] задача решается асимптотически (для достаточно больших  $T$ ) при помощи преобразования Лапласа и введения моменто-производящей функции

$$f(T, \theta) = \sum_{n=1}^{\infty} P_n(T) e^{n\theta}.$$

При больших  $T$  эта функция зависит только от  $g(t)$  [но не от  $w_1(t)$ ] и может быть представлена рядом по степеням  $\theta$ , коэффициенты которого выражаются через (асимптотические) зна-

чения  $\bar{n}$ ,  $\overline{n^2}$ , ... Таким путем получается, в частности, что

$$\bar{n} = \frac{T}{\bar{\vartheta}}, \quad \overline{n^2} - \bar{n}^2 = \frac{\overline{\vartheta^2} - \bar{\vartheta}^2}{\bar{\vartheta}^3} T, \quad \text{где } \overline{\vartheta^m} = \int_0^\infty \vartheta^m g(\vartheta) d\vartheta. \quad (12.6)$$

Для некоррелированных импульсов (12.1) имеет вид (§ 8)

$$\varphi(\vartheta) d\vartheta = n_1^0 d\vartheta,$$

где  $n_1^0$  — среднее число импульсов в единицу времени в этом случае. Соответственно формулы (12.2) и (12.3) дают следующие выражения для вероятности пустого промежутка  $\vartheta$  после импульса:

$$W(\vartheta) = e^{-n_1^0 \vartheta},$$

и для вероятности промежутка  $(\vartheta, \vartheta + d\vartheta)$  между последовательными импульсами:

$$g(\vartheta) d\vartheta = e^{-n_1^0 \vartheta} n_1^0 d\vartheta. \quad (12.7)$$

Если для больших  $\vartheta$  отождествить  $g(\vartheta) d\vartheta$  с  $w_1(\vartheta) d\vartheta$ , т. е. с вероятностью того, что *первый* (он же последний) импульс появился в конце пустого промежутка  $(0, \vartheta)$ , то прямой расчет по формуле (12.5) дает

$$w_n(t) = \frac{e^{-n_1^0 t} (n_1^0)^n t^{n-1}}{(n-1)!},$$

а по (12.4) получаем распределение Пуассона:

$$P_n(T) = \frac{e^{-n_1^0 T} (n_1^0 T)^n}{n!},$$

согласно которому

$$\bar{n} = n_1^0 T, \quad \overline{n^2} - \bar{n}^2 = \bar{n} = n_1^0 T.$$

К тому же результату приводят, конечно, и формулы (12.6) при

$$g(\vartheta) = e^{-n_1^0 \vartheta} n_1^0.$$

Рассмотрим теперь примеры иного задания  $g(\vartheta)$ .

1. Пусть  $g(\vartheta)$  — плотность гамма-распределения:

$$g(\vartheta) = \frac{\kappa^\alpha \vartheta^{\alpha-1}}{\Gamma(\alpha)} e^{-\kappa \vartheta}.$$

Если  $0 < \alpha < 1$ , то вероятность промежутка  $\vartheta$  между импульсами спадает с ростом  $\vartheta$  быстрее, чем в случае  $\alpha = 1$ , который

соответствует отсутствию корреляции, т.е. формуле (12.7) с  $n_1^0 = \kappa$ . Таким образом, при  $0 < \alpha < 1$  мы имеем «притягательную» корреляцию: импульс повышает вероятность последующего импульса. Напротив, при  $\alpha > 1$  корреляция «отталкивательная». Расчет по формулам (12.6) дает

$$\bar{n} = \frac{\kappa T}{\alpha}, \quad \overline{n^2} - \bar{n}^2 = \frac{\kappa T}{\alpha^2} = \frac{\bar{n}}{\alpha}.$$

Из выражения для  $\bar{n}$  видно, что теперь среднее число импульсов в единицу времени есть

$$n_1 = \frac{\kappa}{\alpha} = \frac{n_1^0}{\alpha}.$$

Если в общие выражения (12.6) ввести густоту импульсов  $\kappa$  при отсутствии корреляции, то можно записать их в виде

$$\bar{n} = \gamma \kappa T, \quad \overline{n^2} - \bar{n}^2 = \Gamma^2 \bar{n}, \quad (12.8)$$

где

$$\gamma = \frac{1}{\kappa \bar{\theta}}, \quad \Gamma^2 = \frac{\overline{\theta^2}}{\bar{\theta}^2} - 1. \quad (12.9)$$

В нашем примере

$$\gamma = \frac{1}{\alpha}, \quad \Gamma^2 = \frac{1}{\alpha}.$$

Выражения вида (12.8) получаются при расчете дробового эффекта с учетом влияния объемного заряда. Величина  $\Gamma^2$  называется *коэффициентом депрессии*, так как (в случае  $\Gamma^2 < 1$ ) она показывает, в какой мере дробовой шум подавлен по сравнению с его уровнем при отсутствии объемного заряда. Вычислению  $\Gamma^2$  в зависимости от различных физически интересных условий посвящено много работ. Из теории, развитой Домбом, видно, что  $\gamma$  и  $\Gamma^2$  выражаются в общей форме в терминах корреляции между событиями, так что задача физической теории депрессии дробового эффекта может быть сформулирована как задача отыскания вида распределения  $g(\theta)$  [или  $\varphi(\theta)$ , или  $W(\theta)$ ]<sup>1)</sup>. Подавление дробового шума при наличии объемного заряда означает, что последний создает «отталкивательную» корреляцию между импульсами анодного тока.

<sup>1)</sup> Если вероятность появления импульса зависит от моментов появления не одного, а  $r$  предыдущих импульсов, то формулы (12.8) сохраняются, но для вычисления  $\gamma$  и  $\Gamma^2$  должны быть известны как средние значения  $\bar{\theta}_1$  и  $\bar{\theta}_1^2$ , так и смешанные моменты для интервала  $\theta_1$  и  $r$  последующих интервалов  $\bar{\theta}_1 \bar{\theta}_i$  ( $i = 2, 3, \dots, r$ ). При  $i > r$  длины интервалов уже некоррелированы, так что  $\bar{\theta}_1 \bar{\theta}_i = \bar{\theta}_1^2$ .

2. Счетчик частиц восстанавливается после пролета частицы в течение постоянного промежутка времени  $\beta$ , т. е. вероятность регистрации частицы в течение времени  $\beta$  равна нулю, после чего имеет место равномерное распределение  $\varphi(\theta)d\theta = \lambda d\theta$ . Таким образом,

$$g(\theta) = \begin{cases} 0 & \text{при } \theta < \beta, \\ \lambda e^{-\lambda(\theta-\beta)} & \text{при } \theta > \beta. \end{cases}$$

Это дает

$$\bar{\theta} = \beta + \frac{1}{\lambda}, \quad \overline{\theta^2} - \bar{\theta}^2 = \frac{1}{\lambda^2},$$

так что по (12.9)

$$\gamma = \frac{1}{1 + \lambda\beta}, \quad \Gamma^2 = \frac{1}{(1 + \lambda\beta)^2},$$

т. е. всегда  $\Gamma^2 < 1$ , чего, конечно, и следовало ожидать.

3. Пусть

$$g(\theta) = \beta \delta(\theta) + (1 - \beta) \lambda e^{-\lambda\theta}.$$

Первый член означает, что имеются *конечные* вероятности  $\beta, \beta^2, \dots$  одновременного наступления двух, трех и т. д. событий. В этом случае

$$\bar{\theta} = \frac{1 - \beta}{\lambda}, \quad \overline{\theta^2} - \bar{\theta}^2 = \frac{1 - \beta^2}{\lambda^2}$$

и, следовательно,

$$\gamma = \frac{1}{1 - \beta}, \quad \Gamma^2 = \frac{1 + \beta}{1 - \beta}.$$

Здесь всегда  $\Gamma^2 > 1$ .

Обратимся теперь к случаю *независимых* моментов возникновения импульсов, но с неравномерным распределением  $t_v$ . Вероятность  $n_1(t)dt$  появления импульса в интервале  $(t, t + dt)$  теперь по-прежнему не зависит от того, имелись ли и в каком количестве импульсы до момента времени  $t$ , но зависит от этого времени. Отсчет  $t$  производится от какого-то фиксированного начального момента  $t = 0$ . Ясно, что этим вносится известная упорядоченность: густота толчков будет явно зависеть от  $t$ , и тем самым процесс уже не будет стационарным. Как изменятся формулы для среднего значения, среднего квадрата и функции корреляции импульсного процесса  $\xi(t)$ ?

Можно показать, что и в этом случае для вероятности  $n$  импульсов в интервале  $(0, T)$  остается в силе закон Пуассона с параметром

$$\bar{n} = \int_0^T n_1(t) dt$$

(см. задачу 10). Вероятность же  $w_t(t)dt$  того, что импульс, заведомо появившийся в интервале  $(0, T)$ , возник в промежутке  $(t, t + dt)$  внутри  $(0, T)$ , будет

$$w_t(t)dt = \frac{n_1(t)dt}{\bar{n}}.$$

[При  $n_1(t) = n_1^0 = \text{const}$  мы возвращаемся к распределению  $dt/T$ .] С этими отличиями надо повторить те же выводы, которые были сделаны выше при постоянном  $n_1$  в §§ 10 и 11. Так, например, теперь

$$\begin{aligned} \langle F(t - t_v) F(t + \tau - t_\mu) \rangle &= \\ &= \begin{cases} \frac{1}{\bar{n}^2} \int_{-\infty}^{+\infty} F(\theta) n_1(t - \theta) d\theta \int_{-\infty}^{+\infty} F(\theta' + \tau) n_1(t - \theta') d\theta' & (v \neq \mu), \\ \frac{1}{\bar{n}} \int_{-\infty}^{+\infty} F(\theta) F(\theta + \tau) n_1(t - \theta) d\theta & (v = \mu). \end{cases} \end{aligned}$$

В результате получаем

$$\bar{\xi}(t) = \int_{-\infty}^{+\infty} n_1(t - \theta) \langle F(t, \mathbf{a}) \rangle d\theta, \quad (12.10)$$

$$\begin{aligned} \psi(t, \tau) &= \overline{\xi(t) \xi(t + \tau)} - \bar{\xi}(t) \cdot \bar{\xi}(t + \tau) = \\ &= \int_{-\infty}^{+\infty} n_1(t - \theta) \langle F(\theta, \mathbf{a}) F(\theta + \tau, \mathbf{a}) \rangle d\theta, \end{aligned} \quad (12.11)$$

откуда, в частности,

$$D[\bar{\xi}(t)] = \psi(t, 0) = \overline{\xi^2(t)} - \bar{\xi}(t)^2 = \int_{-\infty}^{+\infty} n_1(t - \theta) \langle F^2(\theta, \mathbf{a}) \rangle d\theta. \quad (12.12)$$

Согласно (12.10) и (12.12) среднее значение и дисперсия  $\bar{\xi}(t)$  теперь не постоянны, а зависят от  $t$ , в чем проявляется нестационарность процесса. Функция корреляции (12.11), характеризующая статистическую связь между  $\xi(t)$  и  $\xi(t + \tau)$ , зависит теперь не только от временного сдвига  $\tau$ , но и от исходного момента времени  $t$ . Разумеется, если положить  $n_1(t) = n_1 = \text{const}$ , то формулы перейдут в полученные ранее<sup>1)</sup>.

Предположим, что  $n_1(t)$  есть некоторая почти-периодическая (в частности, периодическая) функция времени. Тогда такой же почти-периодический (соответственно периодический) характер

<sup>1)</sup> Обобщение на нестационарные импульсные процессы рассматривалось в очень многих работах (см., например, [7]).

будут иметь и моменты (12.10) — (12.12) рассматриваемого нестационарного процесса.

Если на протяжении одного импульса, т. е. за время  $\vartheta$ ,  $n_1(t)$  успевает много раз просциллировать, т. е.  $n_1(t)$  меняется быстро по сравнению с  $F(t)$ , то практически в интегралы даст вклад только среднее значение  $n_1(t - \vartheta)$ , равное  $\bar{n}/T = \bar{n}_1$ . Мы получим тогда

$$\overline{\xi(t)} = \bar{n}_1 \int_{-\infty}^{+\infty} \langle F(\vartheta, a) \rangle d\vartheta, \quad \psi(\tau) = \bar{n}_1 \int_{-\infty}^{+\infty} \langle F(\vartheta, a) F(\vartheta + \tau, a) \rangle d\vartheta,$$

т. е. формулы для стационарного процесса со средним по времени значением числа импульсов в единицу времени  $\bar{n}_1$ .

Такое положение вещей имеет место, например, в усилителе промежуточной частоты. Гетеродин меняет параметры шумящего смесителя с высокой частотой  $\omega_0$ , в результате чего с этой частотой меняется густота  $n_1(t)$  импульсов дробового шума. Длительность же отдельного импульса  $\vartheta$  на выходе усилителя промежуточной частоты определяется, как известно, его полосой  $\Delta\omega \ll \omega_0$  ( $\vartheta \sim 1/\Delta\omega$ ). Таким образом,  $\vartheta \gg 1/\omega_0$  и шум на выходе усилителя промежуточной частоты можно рассматривать как стационарный.

Напротив, если изменение  $n_1(t)$  очень медленное, т. е. на протяжении длительности импульса  $\vartheta$  функция  $n_1(t)$  практически не меняется, то в интегралах можно принять  $n_1(t - \vartheta) = n_1(t)$  и вынести этот множитель из-под интегралов. Мы получаем тогда квазистационарное изменение статистических характеристик процесса:

$$\xi(t) = n_1(t) \int_{-\infty}^{+\infty} \langle F(\vartheta, a) \rangle d\vartheta,$$

$$\psi(t, \tau) = n_1(t) \int_{-\infty}^{+\infty} \langle F(\vartheta, a) F(\vartheta + \tau, a) \rangle d\vartheta.$$

В каждый момент времени  $t$  процесс ведет себя как стационарный со значением густоты импульсов, взятым в этот момент. Так будет обстоять дело, например, при прохождении шума через линейную систему с медленно «ползающими» параметрами.

В дальнейшем нам еще не раз придется вернуться к случайным процессам, представляющим собой хаотическую последовательность импульсов, но уже на основе тех методов, которыми располагает теория случайных функций. Прежде чем перейти к этим методам, остановимся на вопросе о том, как связан асимптотический результат, полученный в § 10, с центральной предельной теоремой теории вероятностей.

### § 13. Импульсный шум и центральная предельная теорема

Как известно, под названием «центральной предельной теоремы» в теории вероятностей фигурирует не одна, а целое семейство теорем различной степени общности, касающихся одного и того же вопроса — предельного распределения суммы независимых случайных величин при неограниченном увеличении числа слагаемых.

Мы видели (§ 9), что сумма *любого* числа нормально распределенных величин тоже распределена по гауссову закону. Центральная предельная теорема сразу же выводит за пределы «нормального семейства». Уже в частном примере пуассоновского импульсного случайного процесса (§ 10) мы убедились, что при увеличении «густоты» шума, т. е. увеличении числа независимых импульсов  $x_v = F(t - t_v, a_v)$ , складывающихся в каждый данный момент времени  $t$ , распределение суммы  $\xi(t) = \sum_v x_v$  стремится к нормальному. Слагаемые  $x_v$  были при этом распределены не по гауссову, но по некоторому специальному закону, *одинаковому* для всех  $x_v$  (и вытекающему из распределений случайных параметров  $a_v$  и  $t_v$ ).

Другой пример того же рода дает теорема Муавра — Лапласа (§ 6), которую нетрудно записать в такой форме, чтобы она относилась к среднему арифметическому (т. е. к сумме независимых случайных величин-фиксаторов  $\epsilon_i$ ):

$$\xi_N = \frac{1}{N} \sum_{i=1}^N \epsilon_i = \frac{n}{N}.$$

Согласно теореме при неограниченном росте числа  $N$  независимых случайных слагаемых  $\epsilon_i$  распределение (нормированной) суммы  $\xi_N$  стремится к нормальному. Здесь все слагаемые распределены, как и в примере импульсного процесса, *одинаково*, но теперь это распределение — простая альтернатива ( $\epsilon_i = 1$  или  $0$  с вероятностями  $p$  и  $q = 1 - p$ ).

Не будет ли нормальное распределение суммы предельным при иных распределениях слагаемых или даже при распределениях, различных для каждого слагаемого? Если это так, то насколько жестким ограничениям должны быть подчинены распределения слагаемых? На эти вопросы и отвечает центральная предельная теорема.

Эти вопросы важны потому, что они касаются статистики таких физических величин, на значение которых влияет очень много независимых случайных факторов, причем каждый из них вызывает лишь весьма малое случайное отклонение. С таким положением вещей мы сталкиваемся и при измерениях, и в



флуктуациях макроскопических величин, обусловленных хаотичностью положений и движений огромного количества микрочастиц — молекул, ионов, электронов и т. д. Эта проблема в свое время привлекла внимание уже Гаусса. Чебышев, Марков, Ляпунов дали первое ее решение, показав, что на случайные слагаемые действительно достаточно наложить лишь очень общие ограничения. Смысл этих ограничений именно в том и состоит, что отдельные слагаемые должны мало влиять на сумму, и тем меньше, чем больше  $N$ .

Строгое доказательство теоремы дал в 1898 г. Марков, а затем в 1900 г. в более общей форме — Ляпунов. Теорема Ляпунова (вместе с дальнейшими ее обобщениями) и получила название *центральной предельной теоремы* теории вероятностей.

Поясним идею доказательства на очень простом случае теоремы *Линдберга — Леви*. В этой теореме (доказанной позже теоремы Ляпунова) содержится очень сильное ограничение: предполагается, что *все слагаемые  $x_v$  распределены одинаково*. Содержание теоремы следующее: пусть  $x_v$  — независимые случайные величины с одинаковой функцией распределения  $W(x)$ , причем  $\bar{x}_v = 0$ ,  $\overline{x_v^2} = 1$ ; тогда, равномерно для всех  $x$ , распределение суммы

$$S_N = \frac{1}{\sqrt{N}} \sum_1^N x_v$$

стремится при  $N \rightarrow \infty$  к

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-x^2/2} dx,$$

т. е.

$$[W(x\sqrt{N})]^{N*} \rightarrow \Phi(x),$$

откуда

$$\left[W\left(\frac{x-a}{\sigma}\right)\right]^{N*} \rightarrow \Phi\left(\frac{x-Na}{\sqrt{N}\sigma}\right)$$

равномерно относительно  $x$  при любых фиксированных  $a$  и  $\sigma$ . Это, конечно, весьма частный случай центральной предельной теоремы, но, в отличие от теоремы Ляпунова, здесь не предполагается существования у  $x_v$  абсолютных моментов порядка  $k > 2$  (см. ниже).

Доказательство теоремы очень простое — благодаря использованию характеристических функций. Пусть  $W(x)$  соответствует характеристическая функция  $\varphi(u)$ . По условию  $\alpha_1 = \bar{x}_v = 0$ ,  $\alpha_2 = \overline{x_v^2} = 1$ . Следовательно,

$$\varphi(u) = 1 - \frac{u^2}{2} + o(u^2),$$

откуда

$$\Phi\left(\frac{u}{\sqrt{N}}\right) = 1 - \frac{u^2}{2N} + o(N^{-3/2}).$$

Для

$$S_N = \frac{1}{\sqrt{N}} \sum_1^N x_v$$

характеристическая функция есть

$$\left[\Phi\left(\frac{u}{\sqrt{N}}\right)\right]^N = \left[1 - \frac{u^2}{2N} + o(N^{-3/2})\right]^N.$$

При  $N \rightarrow \infty$  она стремится равномерно в каждом конечном интервале  $u$  к  $e^{-u^2/2}$ , т. е. к характеристической функции для  $\Phi(x)$ . Это все, что нужно, так как по теореме Леви, если последовательность характеристических функций сходится к  $\phi(u)$ , то последовательность соответствующих функций распределения сходится к  $W(x)$ , соответствующей  $\phi(u)$ .

Теорема Ляпунова является более общей в том отношении, что касается суммы *неодинаково распределенных* независимых случайных величин. И здесь, если предположить, что  $\bar{x}_v = 0$  (что, конечно, несколько не уменьшает общности, а делается просто для удобства), а дисперсии  $\sigma_v^2 = \overline{x_v^2}$  *конечны*, то при определенных условиях закон распределения нормированной суммы

$$S_N = \frac{1}{B_N} \sum_1^N x_v$$

(где  $B_N^2 = \sum_1^N \sigma_v^2$  — дисперсия суммы  $\sum_1^N x_v$ ) стремится при  $N \rightarrow \infty$  к  $\Phi(x)$ . Весь вопрос в этих условиях.

Особенно важен случай, когда при  $N \rightarrow \infty$  имеем  $B_N \rightarrow \infty$ , но  $\sigma_v/B_N \rightarrow 0$ , т. е. каждая отдельная компонента  $\sigma_v^2$  составляет исчезающе малую долю от суммарной дисперсии  $B_N^2$ , неограниченно растущей с увеличением  $N$ . В этом случае является *достаточным* условие, фигурирующее в теореме Ляпунова. Обозначим через  $[C_{kN}]^k$  сумму *абсолютных* моментов  $x_v$   $k$ -го порядка:

$$[C_{kN}]^k = \sum_1^N \overline{|x_v|^k}. \quad (13.1)$$

Условие Ляпунова можно записать тогда в виде

$$\lim_{N \rightarrow \infty} \frac{C_{kN}}{B_N} = 0 \quad \text{при } k > 2, \quad (13.2)$$

т. е. требуется, чтобы у величин  $x_v$  существовали абсолютные моменты *какого-либо* порядка  $k > 2$  и чтобы корень  $k$ -й степени из суммы этих моментов возрастал при  $N \rightarrow \infty$  медленнее, чем квадратный корень из суммы дисперсий. Если условие (13.2) выполнено, то распределение  $F_N(x)$  нормированной суммы  $S_N$  стремится при  $N \rightarrow \infty$  к нормальному закону  $\Phi(x)$ . Пользуясь условием (13.2), Ляпунов не только доказал свою теорему, но и дал оценку величины  $|F_N(x) - \Phi(x)|$ , т. е. нашел, как быстро она стремится к нулю с ростом  $N$  [не медленнее, чем  $C \ln N / \sqrt{N}$ ; Крамер усилил эту оценку: при несколько более жестких условиях модуль разности  $F_N(x)$  и  $\Phi(x)$  меньше  $C/\sqrt{N}$ ].

Заметим, что в том же случае  $B_N \rightarrow \infty$  и  $\sigma_v/B_N \rightarrow 0$  при  $N \rightarrow \infty$  самое общее (*необходимое и достаточное*) условие справедливости центральной предельной теоремы — это условие Линдберга<sup>1)</sup>. Мы ограничимся, однако, формулировкой теоремы при условии Ляпунова, поскольку им вполне можно обойтись в подавляющем большинстве физических задач. Более того, как правило, достаточно пользоваться условием Ляпунова при  $k = 3$ , т. е. брать это условие в виде

$$C_{3N}/B_N \rightarrow 0 \quad \text{при} \quad N \rightarrow \infty. \quad (13.3)$$

Для теории шумов условие Ляпунова означает, что нет больших *редких выбросов* (если только сами эти выбросы не подчинены нормальному закону распределения), что в каждый момент складывается чрезвычайно много равноправных (в смысле их относительной малости) случайных независимых величин. При этих условиях, выполняющихся для широкого класса шумов, распределение шума будет нормальным. Но при одинаковом распределении шумы могут существенно различаться по другим свойствам, в частности по своим спектрам. Для шумов, возникающих от наложения независимых импульсов, эти различия зависят от особенностей отдельных слагаемых, от формы складывающихся импульсов.

Заметим, что проведенный выше вывод условия (10.14), при котором распределение суперпозиции случайных импульсов приближается к нормальному, по сути дела означал применение условия Ляпунова (13.3), причем в данной задаче  $n_1 \neq 0$  и есть число  $N$  случайных величин  $x_v$ , определяющих мгновенное значение суммы  $\sum_v x_v$ .

<sup>1)</sup> См. [8], стр. 73.

## З а д а ч и

1. Логарифмически-нормальным (коротко, логнормальным) распределением случайной величины  $x \geq 0$  называется распределение

$$w(x) dx = \frac{dx}{\sqrt{2\pi} \sigma x} \exp \left\{ -\frac{(\ln x - \lambda)^2}{2\sigma^2} \right\}, \quad (1)$$

где  $\lambda = \langle \ln x \rangle$  и  $\sigma^2 = D[\ln x]$ . Нетрудно видеть, что (1) — это стандартное нормальное распределение нормированной случайной величины  $y = (\ln x - \lambda)/\sigma$ , т. е.

$$w(x) dx = \Phi(y) dy, \quad \Phi(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}. \quad (2)$$

Пользуясь преобразованием от  $x$  к  $y$

$$x = e^{\sigma y + \lambda},$$

найти моменты  $\alpha_k = \overline{x^k}$ .

Р е ш е н и е. Имеем

$$\alpha_k = \overline{x^k} = \int_0^{\infty} x^k w(x) dx = \int_{-\infty}^{+\infty} e^{k(\sigma y + \lambda)} \Phi(y) dy = e^{k\lambda + k^2\sigma^2/2},$$

или, поскольку  $\alpha_1 = e^{\lambda + \sigma^2/2}$ ,

$$\alpha_k = e^{k(k-1)\sigma^2/2} \alpha_1^k.$$

Для дисперсии  $x$  получаем

$$D[x] = \alpha_2 - \alpha_1^2 = (e^{\sigma^2} - 1) \alpha_1^2.$$

Если ввести обозначения  $\lambda = \langle \ln x \rangle \equiv \ln x_0$  и  $e^{\sigma^2/2} = \rho$ , то

$$\alpha_1 = x_0 \rho, \quad D[x] = (\rho^2 - 1) \alpha_1^2, \quad \alpha_k = \rho^{k(k-1)} \alpha_1^k.$$

Логнормальное распределение является предельным для произведения независимых случайных величин  $z_k$  при возрастании числа сомножителей  $n$ .

Если  $x = \prod_{k=1}^n z_k$ , то  $\ln x = \sum_{k=1}^n \ln z_k$ , т. е.  $\ln x$  есть сумма  $n$  независимых случайных величин. По центральной предельной теореме [если, конечно, ее условия выполнены для  $\ln x$  (§ 13)] распределение  $\ln x$  должно стремиться к нормальному.

В силу этой же теоремы должно, в частности, стремиться с ростом  $n$  к нормальному и распределение суммы логнормальных величин  $x_k$ , т. е. величины  $X = \sum_{k=1}^n x_k$ . Однако сходимость распределения величины  $Y = (X - \bar{X})/\sqrt{D[X]}$  к  $\Phi(Y)$  оказывается в случае логнормальных слагаемых очень медленной. Из-за этого получается, что в широкой области значений  $\rho$  и (больших)  $n$  распределение  $Y$ , прежде чем приблизиться с дальнейшим ростом  $n$  к нормальному, лучше аппроксимируется не гауссовым, а логнормальным законом. Соответствующие оценки подробно исследованы в работе [9].

2. Как выражается характеристическая функция случайной величины  $y = ax + b$  через характеристическую функцию случайной величины  $x$ ?

Решение. Имеем

$$\varphi_y(u) = \overline{e^{iuy}} = e^{ibu} \overline{e^{iaux}} = e^{ibu} \varphi_x(au).$$

В частности, для нормированной величины  $y = (x - \bar{x})/\sigma_x$

$$\varphi_y(u) = e^{-i\bar{x}u/\sigma} \varphi_x\left(\frac{u}{\sigma_x}\right).$$

3. Найти характеристическую функцию распределения Коши

$$dW(x) = \frac{dx}{\pi(1+x^2)}.$$

Решение. Для этого распределения формула (9.1) дает

$$\varphi_x(u) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{e^{iux} dx}{1+x^2} = \frac{2}{\pi} \int_0^{\infty} \frac{\cos(ux)}{1+x^2} dx = e^{-|u|}.$$

Точка заострения при  $u = 0$  делает невозможным дифференцирование  $\varphi_x(u)$  в нуле, т. е. вычисление моментов по формуле (9.4).

4. Найти характеристическую функцию случайной величины  $\xi_n = n - pN$  при биномиальном распределении для  $n$  и выполнить предельные переходы к характеристическим функциям распределений Пуассона и Лапласа.

Решение. Имеем

$$\begin{aligned} \varphi_{\xi_n}(u) &= \langle e^{iu(n-pN)} \rangle = e^{-iupN} \sum_{n=0}^N e^{iun} P_N(n) = \\ &= e^{-iupN} \sum_{n=0}^N C_N^n (pe^{iu})^n q^{N-n} = e^{-iupN} (pe^{iu} + q)^N = (pe^{iqu} + qe^{-ipu})^N. \end{aligned}$$

При  $p = \bar{n}/N$  и, соответственно,  $q = 1 - \frac{\bar{n}}{N}$  отсюда следует, что

$$\lim_{N \rightarrow \infty} \varphi_{\xi_n}(u) = \lim_{N \rightarrow \infty} e^{-i\bar{n}u} \left[ 1 + \frac{\bar{n}}{N} (e^{iu} - 1) \right]^N = e^{-i\bar{n}u} e^{\bar{n}(e^{iu} - 1)},$$

т. е. получается характеристическая функция величины  $n - \bar{n}$  при распределении Пуассона для  $n$ .

Для нормированной величины  $x_n = (n - pN)/\sqrt{Npq} = \xi_n/\sqrt{Npq}$  характеристическая функция есть

$$\varphi_{x_n}(u) = \varphi_{\xi_n}\left(\frac{u}{\sqrt{Npq}}\right) = (pe^{iu\sqrt{q/Np}} + qe^{-iu\sqrt{p/Nq}})^N.$$

Разлагая экспоненты в ряды по степеням  $u$ , нетрудно убедиться, что при  $N \rightarrow \infty$  получается  $\varphi_{x_n}(u) = e^{-u^2/2}$ , т. е. характеристическая функция нормального закона распределения для нормированной случайной величины  $x_n$ .

5. Найти характеристическую функцию  $n$ -мерного нормального закона распределения.

Решение. Согласно (7.11) и (9.5)

$$\varphi_{\mathbf{r}}(\mathbf{u}) = \frac{1}{(2\pi)^{n/2} |B|^{1/2}} \int_{-\infty}^{+\infty} \exp \left\{ i\mathbf{u}\mathbf{r} - \sum_{j,k} b_{jk}^{-1} x_j x_k \right\} dx_1 \dots dx_n,$$

где  $u\mathbf{r} = \sum_{k=1}^n u_k x_k$ . Последовательное интегрирование по  $x_1$ , по  $x_2$  и т. д. приводит к результату

$$\varphi_{\mathbf{r}}(u) = \exp \left\{ -\frac{1}{2} \sum_{j,k} b_{jk} u_j u_k \right\}.$$

Это можно записать и иначе, поскольку

$$\sum_{j,k=1}^n b_{jk} u_j u_k = \sum_{j,k=1}^n \overline{x_j x_k} u_j u_k = \overline{\left( \sum_{k=1}^n x_k u_k \right)^2} = \overline{(u\mathbf{r})^2}.$$

Следовательно, для нормального распределения

$$\varphi_{\mathbf{r}}(u) = \langle e^{i u \mathbf{r}} \rangle = e^{-(u\mathbf{r})^2/2}.$$

6. Случайные величины  $x_k$  ( $k = 1, 2, \dots$ ) с  $\bar{x}_k = 0$  распределены нормально. Пользуясь характеристической функцией многомерного нормального распределения

$$\varphi(u) = e^{-S(u)}, \quad S(u) = \frac{1}{2} \sum_{j,k} b_{jk} u_j u_k, \quad b_{jk} = b_{kj} = \overline{x_j x_k},$$

показать, что моменты нечетного порядка равны нулю:

$$\langle x_{i_2} x_{i_2} \dots x_{i_{2n+1}} \rangle = 0, \quad (1)$$

а для моментов четного порядка справедливо выражение

$$\langle x_{i_1} x_{i_2} \dots x_{i_{2n}} \rangle = \frac{1}{2^n n!} \sum b_{i_1 i_2} b_{i_3 i_4} \dots b_{i_{2n-1} i_{2n}}, \quad (2)$$

где  $i_k$  — положительные целые числа, а сумма берется по всем перестановкам индексов  $i_1, \dots, i_{2n}$ .

Решение. Формулы (1) и (2) выводятся по общему правилу:

$$\langle x_{i_1} x_{i_2} \dots x_{i_m} \rangle = \frac{\partial^m \varphi(u)}{i_1^m \partial u_{i_1} \partial u_{i_2} \dots \partial u_{i_m}} \Big|_{u=0}.$$

Число перестановок индексов в (2) равно  $(2n)!$ , но получающиеся при этих перестановках члены частично совпадают. А именно, значение каждого слагаемого не меняется при перестановках сомножителей (таких перестановок  $n!$ ), а кроме того, каждый сомножитель не меняется при перестановке его индексов в силу симметрии  $b_{jk}$  (таких перестановок  $2^n$ ). Таким образом, слагаемые в (2) распадаются на группы по  $2^n n!$  одинаковых слагаемых в каждой. Число различных групп равно, следовательно,

$$\frac{(2n)!}{2^n n!} = (2n-1)!!.$$

Можно поэтому записать (2) и в виде суммы  $(2n-1)!!$  различных слагаемых:

$$\langle x_{i_1} x_{i_2} \dots x_{i_{2n}} \rangle = \sum_{\text{п. п.}} b_{i_1 i_2} b_{i_3 i_4} \dots b_{i_{2n-1} i_{2n}},$$

понимая под  $\sum_{p. п.}$  сумму по перестановкам, дающим разные слагаемые (с учетом симметрии  $b_{jk}$ ). Например,

$$\begin{aligned}\langle x_1 x_2 x_3 x_4 \rangle &= b_{12} b_{34} + b_{13} b_{24} + b_{14} b_{23}, \\ \langle x_1 x_2 x_3 x_4 x_5 x_6 \rangle &= b_{12} (b_{34} b_{56} + b_{35} b_{46} + b_{36} b_{45}) + \\ &+ b_{13} (b_{24} b_{56} + b_{25} b_{46} + b_{26} b_{45}) + b_{14} (b_{23} b_{56} + b_{25} b_{36} + b_{26} b_{35}) + \\ &+ b_{15} (b_{23} b_{46} + b_{24} b_{36} + b_{26} b_{34}) + b_{16} (b_{23} b_{45} + b_{24} b_{35} + b_{25} b_{34}).\end{aligned}$$

В частности, если  $i_1 = i_2 = \dots = i_{2n}$ , получаем

$$\langle x^{2n} \rangle = (2n-1)!! \langle x^2 \rangle^n.$$

Если  $\overline{x_k} \neq 0$ , то полученные выше выражения представляют собой *центральные* моменты  $\langle (x_{i_1} - \overline{x_{i_1}})(x_{i_2} - \overline{x_{i_2}}) \dots (x_{i_m} - \overline{x_{i_m}}) \rangle$ .

7. Показать, что для детерминированной функции  $f(x)$  от гауссовой случайной величины  $x$  справедлива формула

$$\langle f(x) \rangle = \sum_{n=0}^{\infty} f^{(2n)}(\bar{x}) \frac{(\sigma^2/2)^n}{n!}, \quad (1)$$

где  $\sigma^2 = D[x] = \overline{x^2} - \bar{x}^2$ .

Решение. Формула (1) получается в результате усреднения разложения  $f(x)$  в ряд Тейлора по степеням  $x - \bar{x}$ , если воспользоваться значениями центральных моментов  $\langle (x - \bar{x})^k \rangle$  гауссовой величины  $x$  (см. предыдущую задачу).

8. Пусть случайная величина  $x$  распределена по закону

$$w(x) dx = A(a) F(x) e^{af(x)} dx, \quad (1)$$

где  $A(a)$  — нормировочная постоянная. Показать, что для кумулянтов случайной величины  $f(x)$  справедливо выражение

$$\lambda_n(a) = \frac{d^{n-1} \lambda_1(a)}{da^{n-1}}. \quad (2)$$

Решение. Характеристическая функция для  $f(x)$  есть

$$\varphi_f(iu) = \langle e^{iu f(x)} \rangle = A(a) \int F(x) e^{(a+iu)f(x)} dx,$$

где интегрирование распространено на всю область возможных значений  $x$ . Следовательно,

$$\ln \varphi_f(iu) = \ln A(a) + \ln \int F(x) e^{(a+iu)f(x)} dx = \ln A(a) + \mathcal{F}(a+iu).$$

Для первого кумулянта  $f(x)$  имеем

$$\begin{aligned}\lambda_1(a) = \langle f(x) \rangle &= \left. \frac{\partial \ln \varphi_f(iu)}{\partial iu} \right|_{u=0} = \left. \frac{\partial \mathcal{F}(a+iu)}{\partial iu} \right|_{u=0} = \\ &= \left. \frac{\partial \mathcal{F}(a+iu)}{\partial a} \right|_{u=0} = \frac{d\mathcal{F}(a)}{da}.\end{aligned}$$

Таким образом,  $n$ -й кумулянт равен

$$\lambda_n(a) = \frac{\partial^n \ln \varphi_f(iu)}{\partial (iu)^n} \Big|_{u=0} = \left[ \frac{\partial^{n-1}}{\partial a^{n-1}} \frac{\partial \mathcal{F}(a + iu)}{\partial iu} \right]_{u=0} = \frac{d^{n-1} \lambda_1(a)}{da^{n-1}}.$$

Заметим, что (1) — это, в частности, распределение Гиббса для системы с гамильтоновой функцией вида  $H(x) = H_0(x) - bf(x)$ :

$$w(x) dx = A e^{-\beta [H_0(x) - bf(x)]} dx \quad (\beta = 1/kT),$$

т. е. в этом случае  $F(x) = e^{-\beta H_0(x)}$  и  $a = \beta b$ .

9. Найти характеристическую функцию и плотность вероятности случайной величины  $z = f(x)$ , где  $x$  —  $n$ -мерная случайная величина с законом распределения  $w_n(x) dx = w_n(x_1, \dots, x_n) dx_1 \dots dx_n$ , а  $f$  — детерминированная функция.

Решение. Характеристическая функция  $z$  есть

$$\varphi_z(u) = \langle e^{iuz} \rangle = \langle e^{iuf(x)} \rangle = \int e^{iuf(x)} w_n(x) dx.$$

Следовательно,

$$w(z) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iuz} \varphi_z(u) du = \left\langle \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iu[z-f(x)]} du \right\rangle = \langle \delta[z - f(x)] \rangle.$$

Смысл последнего выражения становится вполне ясным, если записать его в развернутом виде:

$$w(z) = \int \delta[z - f(x)] w_n(x) dx.$$

Из всей области возможных значений  $x$  дельта-функция выделяет как раз ту подобласть, в которой  $f(x)$  имеет значение  $z$ . Тем самым  $w(z)dz$  дает вероятностную меру интервала  $(z, z + dz)$  для  $z$ .

10. Показать, что для усредненного дробового тока (10.10) дисперсия выражается формулой (5.2), если время усреднения  $T$  заметно превышает длительность импульса  $\theta$ .

Решение. В соответствии с (10.10)

$$\begin{aligned} D[I_T] &= \overline{(I_T - \bar{I}_T)^2} = \frac{1}{T^2} \int_0^T \int_0^T [\overline{I(t) I(t')} - \bar{I}(t) \cdot \bar{I}(t')] dt dt' = \\ &= \frac{1}{T^2} \iint_0^T \psi_I(\tau) dt dt', \end{aligned}$$

где  $\tau = t - t'$ . Так как корреляционная функция  $\psi_I(\tau)$  — четная функция  $\tau$ , можно написать

$$D[I_T] = \frac{2}{T^2} \int_0^T dt \int_t^T dt' \psi_I(t - t') = \frac{2}{T^2} \int_0^T dt \int_0^{T-t} \psi_I(\tau) d\tau.$$



Функция корреляции мгновенного тока  $I(t)$  выражается формулой (11.6), если в ней положить  $a = e$  и считать площадь импульса единичной:

$$\psi_I(\tau) = n_1 e^2 \int_{-\infty}^{+\infty} F(\theta) F(\theta + \tau) d\theta, \quad \int_{-\infty}^{+\infty} F(\theta) d\theta = 1.$$

Достаточно вычислить  $D[I_T]$  для какой-либо простейшей формы импульса, например для прямоугольного импульса (рис. 3). В этом случае получается «треугольная» функция корреляции:

$$\psi_I(\tau) = \begin{cases} \frac{e\bar{I}}{\Phi} \left(1 - \frac{|\tau|}{\Phi}\right) & \text{при } |\tau| < \Phi, \\ 0 & \text{при } |\tau| \geq \Phi. \end{cases}$$

Подставив ее в двукратный интеграл для  $D[I_T]$ , находим

$$D[I_T] = \begin{cases} \frac{e\bar{I}}{T} \left(1 - \frac{\Phi}{3T}\right) & \text{при } T > \Phi, \\ \frac{e\bar{I}}{\Phi} \left(1 - \frac{T}{3\Phi}\right) & \text{при } T < \Phi. \end{cases}$$

Таким образом, с увеличением времени усреднения  $T$  устанавливается значение дисперсии (5.2). Практически достаточно уже небольшого превышения  $T$  над длительностью импульса  $\Phi$ .

11. При помощи вероятностей  $P_{\text{чет}}$  и  $P_{\text{неч}}$ , вычисленных в § 11, рассчитать двумерные характеристическую функцию и плотность вероятности для случайного телеграфного сигнала.

Решение. Если в момент времени  $t$  значения  $\pm a$  имеют вероятности  $1/2$ , то характеристическая функция двумерного распределения будет (полагая  $a = 1$ )

$$\begin{aligned} \varphi_2(u_1, u_2, \tau) &= e^{i(u_1 \xi + u_2 \xi_\tau)} = \\ &= e^{i(u_1 + u_2)} \frac{P_{\text{чет}}}{2} + e^{-i(u_1 + u_2)} \frac{P_{\text{чет}}}{2} + e^{i(u_1 - u_2)} \frac{P_{\text{неч}}}{2} + e^{-i(u_1 - u_2)} \frac{P_{\text{неч}}}{2} = \\ &= e^{-\bar{n}} [\operatorname{ch} \bar{n} \cdot \cos(u_1 + u_2) + \operatorname{sh} \bar{n} \cdot \cos(u_1 - u_2)]. \end{aligned}$$

Отсюда, обращением по Фурье, находим

$$\begin{aligned} w_2(\xi, \xi_\tau, \tau) &= \frac{1}{4\pi^2} \iint_{-\infty}^{+\infty} e^{-i(u_1 \xi + u_2 \xi_\tau)} \varphi_2(u_1, u_2, \tau) du_1 du_2 = \\ &= \frac{1 + e^{-2\bar{n}}}{4} [\delta(\xi - 1) \delta(\xi_\tau - 1) + \delta(\xi + 1) \delta(\xi_\tau + 1)] + \\ &\quad + \frac{1 - e^{-2\bar{n}}}{4} [\delta(\xi - 1) \delta(\xi_\tau + 1) + \delta(\xi + 1) \delta(\xi_\tau - 1)], \end{aligned}$$

причем  $\bar{n} = n_1 \tau$ .

12. В литературе был описан класс двумерных эллиптически-симметричных распределений [10], обобщающих двумерный гауссов закон для величин  $x_1$  и  $x_2$ , обладающих одинаковыми дисперсиями. При  $\sigma_1^2 = \sigma_2^2 = \sigma^2$  гауссова

плотность вероятности (7.13) имеет вид

$$w_2(x_1, x_2) = \frac{1}{2\pi\sigma^2 \sqrt{1-K^2}} \exp \left\{ -\frac{r^2}{2\sigma^2(1-K^2)} \right\}, \quad (1)$$

т. е. зависит от  $x_1$  и  $x_2$  только через

$$r = \sqrt{x_1^2 + x_2^2 - 2Kx_1x_2}.$$

У эллиптически-симметричных распределений  $w_2$  тоже зависит только от  $r$ :

$$w_2(x_1, x_2) = F(r),$$

но  $F$  — любая неотрицательная и нормированная к единице функция. Такие распределения, как и их частный случай (1), удобны для исследования ряда вопросов, касающихся прохождения сигналов и шумов через безынерционные нелинейные устройства.

Показать, что двумерная характеристическая функция для эллиптически-симметричных распределений может быть записана в виде модифицированного преобразования Бесселя от  $F(r)$ :

$$\varphi_2(u_1, u_2) \equiv \bar{\varphi}(s) = \frac{2\pi}{\sqrt{1-K^2}} \int_0^\infty F(r) J_0\left(\frac{sr}{\sqrt{1-K^2}}\right) r dr, \quad (2)$$

где  $J_0$  — нулевая бesselова функция и

$$s = \sqrt{u_1^2 + u_2^2 + 2Ku_1u_2}.$$

Найти одномерную характеристическую функцию.

Решение. По определению

$$\varphi_2(u_1, u_2) = \iint_{-\infty}^{+\infty} F(r) \exp \{i(u_1x_1 + u_2x_2)\} dx_1 dx_2.$$

Перейдя на плоскости  $(x_1, x_2)$  к координатам  $r$  и  $\alpha$ :

$$x_1 - Kx_2 = r \cos \alpha, \quad x_2 \sqrt{1-K^2} = r \sin \alpha,$$

$$dx_1 dx_2 = \frac{r dr d\alpha}{\sqrt{1-K^2}},$$

получаем

$$\varphi_2(u_1, u_2) = \frac{1}{\sqrt{1-K^2}} \int_0^\infty F(r) r dr \int_0^{2\pi} d\alpha \exp \left\{ ir \left[ u_1 \cos \alpha + \frac{Ku_1 + u_2}{\sqrt{1-K^2}} \sin \alpha \right] \right\}.$$

Как известно,

$$\int_0^{2\pi} d\alpha \exp \{i[a \cos \alpha + b \sin \alpha]\} = 2\pi J_0(\sqrt{a^2 + b^2}).$$

Пользуясь этой формулой, получаем результат (2).

Из определения характеристической функции следует, что одномерная функция  $\varphi_1(u_1)$  получается из двумерной  $\varphi_2(u_1, u_2)$ , если положить в последней  $u_2 = 0$ :

$$\varphi_1(u_1) = \varphi_2(u_1, 0).$$

Поэтому из формулы (2) следует, что для эллиптически-симметричных распределений

$$\varphi_1(u_1) = \frac{2\pi}{\sqrt{1-K^2}} \int_0^\infty F(r) J_0\left(\frac{u_1 r}{\sqrt{1-K^2}}\right) r dr = \tilde{\varphi}(u_1),$$

т. е. одномерная характеристическая функция зависит от своего аргумента так же, как двумерная от  $s$ :

$$\varphi_2(u_1, u_2) = \tilde{\varphi}(s) = \varphi_1(s).$$

Таким образом, задавая четную одномерную плотность вероятности  $w_1(x)$  и находя соответствующую четную по  $s$  характеристическую функцию

$$\varphi_1(s) = \int_{-\infty}^{+\infty} e^{isx} w_1(x) dx = 2 \int_0^\infty w_1(x) \cos sx \cdot dx,$$

можно посредством преобразования, обратного (2), получить затем двумерную плотность вероятности  $w_2(x_1, x_2) = F(r)$ :

$$F(r) = \frac{1}{2\pi \sqrt{1-K^2}} \int_0^\infty \varphi_1(s) J_0\left(\frac{sr}{\sqrt{1-K^2}}\right) s ds.$$

Следует, однако, иметь в виду, что пригодны не любые четные  $w_1(x)$ , так как не все они приводят к неотрицательной и интегрируемой  $w_2(x_1, x_2)$ .

13. Показать, что из экспоненциального распределения (5.4) для времени ожидания события в хаотическом потоке событий вытекает распределение Пуассона (5.3) для вероятности наступления  $n$  событий в фиксированном интервале времени  $(0, T)$ .

Решение. Пусть  $t_1, t_2, \dots$  — моменты наступления событий 1, 2, ..., отсчитываемые от  $t = 0$ . Согласно (5.4) вероятность того, что событие 1 произойдет в  $(t_1, t_1 + dt_1)$ , есть

$$u(t_1) dt_1 = e^{-n_1 t_1} n_1 dt_1, \quad \text{где } 0 \leq t_1 < \infty.$$

Вероятность наступления события 2 в  $(t_2, t_2 + dt_2)$ , т. е. спустя время  $\tau_2 = t_2 - t_1$  после события 1, есть

$$u(\tau_2) dt_2 = e^{-n_1 \tau_2} n_1 dt_2, \quad \text{где } t_1 \leq t_2 < \infty,$$

и т. д., вплоть до вероятности наступления  $(n+1)$ -го события в интервале  $(t_{n+1}, t_{n+1} + dt_{n+1})$ , равной

$$u(\tau_{n+1}) dt_{n+1} = e^{-n_1 \tau_{n+1}} n_1 dt_{n+1}, \quad \text{где } \tau_{n+1} = t_{n+1} - t_n \text{ и } t_n \leq t_{n+1} < \infty.$$

Так как промежутки времени  $\tau_k$  между событиями независимы, совместная вероятность наступления  $n+1$  событий в моменты  $t_1 \leq t_2 \leq \dots \leq t_{n+1}$  равна

$$\begin{aligned} w_{n+1}(\tau_1, \dots, \tau_{n+1}) dt_1 \dots dt_{n+1} &= \prod_{k=1}^{n+1} u(\tau_k) dt_k = \\ &= e^{-n_1(\tau_1 + \dots + \tau_{n+1})} n_1^{n+1} dt_1 \dots dt_{n+1} = e^{-n_1 t_{n+1}} n_1^{n+1} dt_1 \dots dt_{n+1}. \end{aligned}$$

Нас интересует вероятность  $P(n)$  наступления  $n$  событий в интервале  $(0, T)$ , т. е.  $n$ -е событие происходит еще внутри  $(0, T)$ , но  $(n+1)$ -е — уже

вне  $(0, T)$ . Поэтому  $P(n)$  равно интегралу от  $\omega_{n+1}$  по  $(n+1)$ -мерной области, выделенной условиями  $0 \leq t_1 \leq t_2 \leq \dots \leq t_n \leq T \leq t_{n+1} < \infty$ :

$$\begin{aligned} P(n) &= \int_T^\infty dt_{n+1} \int_0^T dt_n \int_0^{t_n} dt_{n-1} \dots \int_0^{t_2} \omega_{n+1} dt_1 = \\ &= \int_T^\infty e^{-n_1 t_{n+1}} n_1 dt_{n+1} \cdot n_1^n \int_0^T dt_n \int_0^{t_n} dt_{n-1} \dots \int_0^{t_2} dt_1. \end{aligned}$$

Первый интеграл (по  $t_{n+1}$ ) равен  $e^{-n_1 T}$  и представляет собой вероятность того, что  $(n+1)$ -е событие произошло после момента  $T$ . Остальные интегралы (после знака умножения) дают  $\frac{(n_1 T)^n}{n!}$ , и это есть вероятность того, что  $n$  предыдущих событий наступили в  $(0, T)$ . В результате получаем

$$P(n) = \frac{\bar{n}^n}{n!} e^{-\bar{n}}, \quad \text{где } \bar{n} = n_1 T,$$

т. е. закон Пуассона (5.3).

14. Показать, что для потока независимых событий с неравномерным распределением  $n_1(t) dt$  вероятности наступления события в интервале  $(t, t+dt)$

остается в силе закон Пуассона с параметром  $\bar{n} = \int_0^T n_1(t) dt$ .

Решение. Вероятность пустого интервала  $(0, t)$  равна теперь  $\exp \left\{ - \int_0^t n_1(t) dt \right\}$ , так что вероятность ожидания  $n$ -го события в интервале  $(t_n, t_n + dt_n)$  после наступления  $(n-1)$ -го события в момент  $t_{n-1}$  будет

$$u(t_n, t_{n-1}) dt_n = \exp \left\{ - \int_{t_{n-1}}^{t_n} n_1(t) dt \right\}, \quad \text{где } t_{n-1} \leq t_n < \infty.$$

Совместная вероятность наступления  $n+1$  событий в моменты времени  $t_1, t_2, \dots, t_{n+1}$ , в силу их независимости, равна ( $t_0 = 0$ )

$$\begin{aligned} \omega_{n+1}(t_1, \dots, t_{n+1}) dt_1 \dots dt_{n+1} &= \prod_{k=1}^{n+1} u(t_k, t_{k-1}) dt_k = \\ &= \exp \left\{ - \sum_{k=1}^{n+1} \int_{t_{k-1}}^{t_k} n_1(t) dt \right\} \cdot \prod_{k=1}^{n+1} n_1(t_k) dt_k = \\ &= \exp \left\{ - \int_0^{t_{n+1}} n_1(t) dt \right\} \cdot \prod_{k=1}^{n+1} n_1(t_k) dt_k. \end{aligned}$$

Те же рассуждения, что и в предыдущей задаче, приводят к следующему выражению для вероятности  $n$  событий в интервале  $(0, T)$ :

$$P(n) = \int_0^T \exp \left\{ - \int_0^{t_{n+1}} n_1(t) dt \right\} n_1(t_{n+1}) dt_{n+1} \int_0^T n_1(t_n) dt_n \times \\ \times \int_0^{t_n} n_1(t_{n-1}) dt_{n-1} \dots \int_0^{t_2} n_1(t_1) dt_1 = \frac{\bar{n}}{n!} e^{-\bar{n}},$$

где

$$\bar{n} = \int_0^T n_1(t) dt.$$

15. Предполагая, что логарифм характеристической функции одномерной случайной величины  $x$ , определенной на интервале возможных значений  $(-\infty, \infty)$ , разлагается в бесконечный ряд:

$$\ln \varphi(u) = \sum_{m=1}^{\infty} \frac{\lambda_m}{m!} (iu)^m,$$

показать, что для плотности вероятности  $w(x)$  справедливо соотношение

$$\frac{\partial w(x)}{\partial \lambda_k} = \frac{(-1)^k}{k!} \frac{\partial^k w(x)}{\partial x^k}. \quad (1)$$

Решение. Запишем характеристическую функцию в виде

$$\varphi(u) = \exp \{ \ln \varphi(u) \} = \exp \left\{ \sum_{m=1}^{\infty} \frac{\lambda_m}{m!} (iu)^m \right\}.$$

Отсюда, если кумулянт  $\lambda_k$  отличен от нуля, находим

$$\frac{\partial \varphi(u)}{\partial \lambda_k} = \frac{(iu)^k}{k!} \varphi(u),$$

и поэтому дифференцирование (9.8) по  $\lambda_k$  дает

$$\frac{\partial w(x)}{\partial \lambda_k} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iux} \frac{\partial \varphi(u)}{\partial \lambda_k} du = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{(iu)^k}{k!} e^{-iux} \varphi(u) du.$$

С другой стороны, согласно (9.8)

$$\frac{\partial^k w(x)}{\partial x^k} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} (-iu)^k e^{-iux} \varphi(u) du.$$

Сопоставляя две последние формулы, получаем (1).

16. Показать при помощи соотношения (1) предыдущей задачи:

$$\frac{\partial w(x)}{\partial \lambda_k} = \frac{(-1)^k}{k!} \frac{\partial^k w(x)}{\partial x^k}. \quad (1)$$

что для любой функции  $f(x)$  справедливо уравнение [11]

$$\frac{\partial^n \langle f(x) \rangle}{\partial \lambda_k^n} = \frac{1}{(k!)^n} \left\langle \frac{d^{nk} f(x)}{dx^{nk}} \right\rangle. \quad (2)$$

Решение. Дифференцируя равенство

$$\langle f(x) \rangle = \int_{-\infty}^{+\infty} f(x) w(x) dx$$

по  $\lambda_k$  и пользуясь (1), получаем

$$\frac{\partial \langle f(x) \rangle}{\partial \lambda_k} = \int_{-\infty}^{+\infty} f(x) \frac{\partial w(x)}{\partial \lambda_k} dx = \frac{(-1)^k}{k!} \int_{-\infty}^{+\infty} f(x) \frac{\partial^k w(x)}{\partial x^k} dx.$$

$k$ -кратное интегрирование по частям, если учесть, что  $w(x)$  вместе со своими производными обращается в нуль при  $|x| \rightarrow \infty$ , приводит к соотношению

$$\frac{\partial \langle f(x) \rangle}{\partial \lambda_k} = \frac{1}{k!} \left\langle \frac{d^k f(x)}{dx^k} \right\rangle. \quad (3)$$

Применение этой формулы последовательно  $n$  раз приводит к (2). Если продифференцировать (2) еще  $m$  раз по *другому* кумулянту  $\lambda_l$ , то, пользуясь повторно формулой (2), получим

$$\frac{\partial^{n+m} \langle f(x) \rangle}{\partial \lambda_k^n \partial \lambda_l^m} = \frac{1}{(k!)^n (l!)^m} \left\langle \frac{d^{nk+ml} f(x)}{dx^{nk+ml}} \right\rangle. \quad (4)$$

Одно из применений «кумулянтных уравнений» (1)–(4) — быстрое вычисление коэффициентов в разложениях моментов  $\alpha_k = \langle x^k \rangle$  случайной величины  $x$  по кумулянтам  $\lambda_m$  этой величины [11]. Например, в таком разложении  $\alpha_4$  коэффициент при  $\lambda_1 \lambda_3$  равен, согласно (4),

$$\frac{\partial^2 \langle x^4 \rangle}{\partial \lambda_1 \partial \lambda_3} = \frac{1}{1! 3!} \left\langle \frac{d^4 x^4}{dx^4} \right\rangle = \frac{4!}{3!} = 4,$$

а при  $\lambda_1^2 \lambda_2$  —

$$\frac{1}{2} \frac{\partial^3 \langle x^4 \rangle}{\partial \lambda_1^2 \partial \lambda_2} = \frac{1}{2 (1!)^2 2!} \left\langle \frac{d^4 x^4}{dx^4} \right\rangle = \frac{4!}{2 \cdot 2!} = 6.$$

Вывод кумулянтных уравнений, проведенный в двух предыдущих задачах для одномерной случайной величины, легко обобщается на многомерные величины. Пользуясь характеристической функцией двумерной величины  $\xi = \{x, y\}$  [см. (9.5)]

$$\Phi(u, v) = \int_{-\infty}^{+\infty} e^{i(ux+vy)} w(x, y) dx dy$$

и разложением в ряд ее логарифма

$$\ln \Phi(u, v) = \sum_{m+n \geq 1} \frac{(iu)^m (iv)^n}{m! n!} \lambda_{m+n},$$

где  $m$  относится к  $x$ , а  $n$  — к  $y$  ( $m + n$  — порядок кумулянта), нетрудно показать, что для совместной плотности вероятности  $w(x, y)$  справедливо соотношение

$$\frac{\partial w(x, y)}{\partial \lambda_{p+q}} = \frac{(-1)^{p+q}}{p! q!} \left\langle \frac{\partial^{p+q} w(x, y)}{\partial x^p \partial y^q} \right\rangle. \quad (5)$$

Можно показать, далее, что для произвольной функции  $f(x, y)$

$$\frac{\partial \langle f(x, y) \rangle}{\partial \lambda_{p+q}} = \frac{1}{p! q!} \left\langle \frac{\partial^{p+q} f(x, y)}{\partial x^p \partial y^q} \right\rangle. \quad (6)$$

Повторное применение этой формулы  $m$  раз дает, очевидно,

$$\frac{\partial^m \langle f(x, y) \rangle}{\partial \lambda_{p+q}^m} = \frac{1}{(p! q!)^m} \left\langle \frac{\partial^{m(p+q)} f(x, y)}{\partial x^{mp} \partial y^{mq}} \right\rangle, \quad (7)$$

а дифференцирование (7)  $n$  раз по *другому* кумулянту  $\lambda_{p_1+q_1}$  приводит к уравнению

$$\frac{\partial^{m+n} \langle f(x, y) \rangle}{\partial \lambda_{p+q}^m \partial \lambda_{p_1+q_1}^n} = \frac{1}{(p! q!)^m (p_1! q_1!)^n} \left\langle \frac{\partial^{m(p+q)+n(p_1+q_1)} f(x, y)}{\partial x^{mp+np_1} \partial y^{mq+nq_1}} \right\rangle. \quad (8)$$

Кумулянтные уравнения, в частности уравнения (7), (8), очень полезны при рассмотрении нелинейных преобразований случайных величин [11]. В дальнейшем (задача 3 гл. VII) будет приведен пример их использования.

## Г л а в а III

### СЛУЧАЙНЫЕ ФУНКЦИИ

#### § 14. Общие определения

Наглядное представление о случайном процессе можно почерпнуть из самых различных областей физики и техники. Осциллограмма дробового тока электронной лампы или тепловых флуктуаций напряжения на каком-либо сопротивлении, перемещение брауновской частицы, записанные сейсмографом колебания почвы, пульсации давления или температуры в атмосфере — все это примеры случайных процессов. Осуществленная запись наблюдаемой величины, т. е. временная развертка значений, *фактически принятых* ею в данном эксперименте, изображает одну из возможных *реализаций* рассматриваемого процесса.

Если представить себе множество идентичных экземпляров рассматриваемой системы, поставленных в одни и те же условия (*ансамбль* систем), то при одинаковом способе регистрации процесса полученные реализации будут, вообще говоря, различны, в чем и проявляется случайный характер процесса<sup>1)</sup>. Элемент случайности в совокупности реализаций может быть различным. Когда он полностью отсутствует, мы приходим к частному случаю вполне детерминированного течения процесса.

В широком, хотя и не особенно точном понимании случайным процессом  $\xi(t)$  является любой протекающий во времени процесс, управляемый вероятностными законами. Более определено: случайная функция  $\xi(t)$  — это такая функция, значение которой при любом возможном  $t$  есть случайная величина.

В математической теории нет никаких оснований к тому, чтобы придавать аргументу  $t$  какую-либо определенную интер-

---

<sup>1)</sup> Естественно, сразу же возникает вопрос о том, что физически означает *идентичность* систем, образующих ансамбль, но мы обратимся к нему позднее (§ 20).



претацию. Параметр  $t$ , вещественный или комплексный, может быть как непрерывным, так и дискретным. Но в приложениях речь идет большей частью о случайных функциях *времени* и (или) *пространственных координат*. Нас будет интересовать поэтому только тот случай, когда возможные значения  $t$  вещественны. Что касается дискретности или непрерывности значений  $t$ , то могут встретиться обе эти возможности. Например, можно фиксировать координату брауновской частицы  $\xi$  только в отдельные моменты времени  $t_1, t_2, \dots$ , получая тем самым последовательность случайных величин  $\xi(t_1), \xi(t_2), \dots$ . В дальнейшем мы в большинстве случаев будем понимать под  $t$  время. Возможные значения функции  $\xi(t)$  при каком-либо фиксированном  $t = t_k$  будут обозначаться через  $x_k$ .

Принято называть  $\xi(t)$  *случайным* (или стохастическим, или вероятностным) *процессом*, если  $t$  меняется *непрерывно*, и *случайной последовательностью* (процессом с дискретным параметром, с дискретным временем), если  $t$  принимает *счетное* множество значений. Термин «случайная функция» охватывает оба эти случая, т. е. применяется тогда, когда характер изменения  $t$  произволен.

В свою очередь множество возможных значений  $x$  самой случайной функции  $\xi(t)$  тоже может быть как непрерывным, так и дискретным. Таким образом, имеются четыре разновидности случайной функции в соответствии с четырьмя комбинациями дискретности и непрерывности  $x$  и  $t$ . Иногда все эти разновидности называют случайными процессами, говоря о дискретном или непрерывном процессе (в зависимости от характера  $x$ ) с дискретным или непрерывным временем  $t$ . Разумеется, возможны и смешанные — дискретно-непрерывные — процессы, но изучать удобнее каждый из этих видов в отдельности.

Если аргумент  $t$  принимает *конечное* множество значений  $t_1, t_2, \dots, t_n$ , то случайная последовательность сводится, очевидно, к совокупности  $n$  случайных величин  $\xi(t_1), \xi(t_2), \dots, \xi(t_n)$ , т. е. может быть интерпретирована как  $n$ -мерная случайная величина, задаваемая, как обычно, своим  $n$ -мерным распределением вероятности  $\omega_n(x_1, x_2, \dots, x_n)$ . Если же множество значений  $t$  бесконечно (счетно или непрерывно), то мы выходим за рамки классической теории вероятностей и необходимо специально определить, как в этом случае надо понимать *задание* случайной функции  $\xi(t)$ .

Беря за исходный пункт множество всех возможных реализаций случайной функции, можно получить исчерпывающую ее характеристику заданием распределения вероятностей этих реализаций. Равносильный, но по форме отличный подход, принадлежащий Е. Е. Слуцкому, опирается на то, что  $\xi(t)$  при каждом фиксированном значении  $t$  есть случайная величина. Эта

величина полностью задана, если известно ее распределение

$$\omega_1(t, x) dx = P\{x < \xi(t) \leq x + dx\},$$

т. е. известна мгновенная плотность вероятности  $\omega_1(t, x)$ , вообще говоря, зависящая от  $t$ . Разумеется, вместо  $\omega_1(t, x)$  можно задать соответствующую характеристическую функцию

$$\varphi_1(t, u) = \int_{-\infty}^{+\infty} e^{iux} \omega_1(t, x) dx.$$

Легко видеть, однако, что, зная распределение  $\xi(t)$  для любого  $t$ , мы располагаем еще в высшей степени неполной характеристикой случайной функции. Возможны функции  $\xi(t)$ , обладающие одинаковыми распределениями  $\omega_1(t, x)$ , но различающиеся статистическим соотношением между значениями  $\xi(t_1)$  и  $\xi(t_2)$ , принимаемыми в два разных момента времени  $t_1$  и  $t_2$ . Другими словами, при одинаковых распределениях  $\omega_1(t, x)$  у этих функций могут быть различны двумерные распределения  $\omega_2(t_1, x_1; t_2, x_2) dx_1 dx_2 =$

$$= P\{x_1 < \xi(t_1) \leq x_1 + dx_1; x_2 < \xi(t_2) \leq x_2 + dx_2\}$$

и, в частности, могут быть различны функции корреляции

$$\psi(t_1, t_2) = \overline{\xi(t_1) \xi(t_2)} - \overline{\xi(t_1)} \overline{\xi(t_2)}.$$

Разумеется,

$$\omega_1(t, x) = \int_{-\infty}^{+\infty} \omega_2(t, x; t', x') dx' = \int_{-\infty}^{+\infty} \omega_2(t', x'; t, x) dx',$$

т. е., зная двумерное распределение, мы знаем и одномерное, но не наоборот, так что совпадение одномерных распределений для каких-либо случайных функций еще ничего не говорит об их двумерных распределениях.

В свою очередь задание двумерного распределения, т. е. плотности вероятности  $\omega_2(t_1, x_1; t_2, x_2)$  или характеристической функции

$$\varphi_2(t_1, u_1; t_2, u_2) = \int_{-\infty}^{+\infty} e^{i(u_1 x_1 + u_2 x_2)} \omega_2(t_1, x_1; t_2, x_2) dx_1 dx_2,$$

которые в общем случае зависят от двух параметров  $t_1$  и  $t_2$ , давая более полную характеристику  $\xi(t)$ , тоже не является исчерпывающим. Двумерное распределение не позволяет судить о статистических соотношениях между значениями  $\xi(t)$ , принимаемыми в какие-либо три момента времени  $t_1$ ,  $t_2$  и  $t_3$ . Эти соотношения могут быть различны при одинаковом двумерном (а зна-

чит, и одномерном) распределении. Ясно, что это рассуждение можно продолжить, и оно приведет тогда к тому, что для любого фиксированного числа моментов времени  $t_i$  ( $i = 1, 2, \dots, n$ )  $n$ -мерное распределение

$$\omega_n(t_1, x_1; \dots; t_n, x_n) dx_1 \dots dx_n = \\ = P\{x_i < \xi(t_i) \leq x_i + dx_i; i = 1, 2, \dots, n\}$$

дает полные сведения о  $\xi(t)$  «вниз» от  $n$ , но оставляет столь же полную неопределенность «вверх». В этом и состоит принципиальное отличие случайной функции, у которой аргумент  $t$  может принимать бесконечное множество значений, от частного случая, когда это множество конечно ( $n$  значений) и дело сводится к  $n$ -мерной случайной величине.

Для практических целей можно было бы ограничиться заданием  $n$ -мерного распределения с весьма большим  $n$ , но подобный выход из положения был бы столь же мало удовлетворительным, как если бы, например, мы согласились рассматривать в анализе вместо непрерывно меняющихся величин только дискретные изменения, хотя бы и с достаточно малым шагом. Можно, однако, считать, что нам известны *конечномерные* ( $n$ -мерные) распределения, но для *всякого* сколь угодно *большого* значения  $n$ . Мы приходим, таким образом, к следующему исчерпывающему способу задания случайной функции: *случайная функция задана, если ее конечномерное распределение  $\omega_n(t_1, x_1; \dots; t_n, x_n)$  известно для любого числа  $n$  произвольно выбранных значений  $t_1, t_2, \dots, t_n$ .*

Очевидно,  $\omega_n(t_1, x_1; \dots; t_n, x_n)$  должны быть *симметричны* относительно любых перестановок всех пар аргументов  $(t_i, x_i)$ , так как содержание рассматриваемого события — совместное, осуществление  $n$  неравенств  $x_i < \xi(t_i) \leq x_i + dx_i$  — не зависит от того, в каком порядке эти неравенства перечислять. Например,

$$P\{x_i < \xi(t_i) \leq x_i + dx_i; i = 1, 2\} = \\ = \omega_2(t_1, x_1; t_2, x_2) dx_1 dx_2 = \omega_2(t_2, x_2; t_1, x_1) dx_1 dx_2.$$

Считая это условие симметрии выполненным, мы часто не будем писать всех аргументов функций  $\omega_n$ .

Кроме условия симметрии, все конечномерные плотности вероятности должны быть еще *согласованы* между собой в смысле их соподчинения, т. е. любое  $k$ -мерное распределение должно определяться из всякого  $n$ -мерного с  $n > k$ :

$$\omega_k(t_1, x_1; \dots; t_k, x_k) = \\ = \int_{-\infty}^{+\infty} \omega_n(t_1, x_1; \dots; t_k, x_k; \dots; t_n, x_n) dx_{k+1} \dots dx_n.$$

В рассматриваемой теории случайных *функций*, когда  $w_n$  зависят от параметров  $t_i$  ( $i = 1, \dots, n$ ), это условие, ограничивающее класс допустимых  $w_n$ , далеко не тривиально. В самом деле, интегрирование  $w_n$  по «лишним» переменным  $x_{k+1}, \dots, x_n$  должно автоматически приводить к результату, не зависящему от параметров  $t_{k+1}, \dots, t_n$ , содержащихся в подынтегральной функции.

Условия симметрии и согласованности должны выполняться для *всякой* случайной функции. Дальнейшая детализация свойств конечномерных распределений ведет уже к установлению специальных классов или типов этих функций. Имеется сравнительно немного таких классов, практическое значение которых особенно велико и для которых теория продвинута в наибольшей степени. Среди них должны быть в первую очередь названы *случайные процессы марковского типа* (иначе — процессы без вероятностного последствия) и *стационарные процессы*<sup>1)</sup>. Мы приведем сначала определения обоих этих классов случайных процессов, после чего остановимся на каждом из них более подробно. Подчеркнем сразу же, что эти классы выделены по *разным* признакам, т. е. их границы не совпадают. Будучи марковским, случайный процесс может являться как стационарным, так и нестационарным; будучи стационарным, он может быть или не быть марковским.

## § 15. Марковские процессы

Выберем  $n$  последовательных моментов времени  $t_1 < \dots < t_n$ . По определению условной вероятности мы всегда можем написать

$$\begin{aligned} w_n(t_1, x_1; \dots; t_n, x_n) dx_1 \dots dx_n = \\ = w_{n-1}(t_1, x_1; \dots; t_{n-1}, x_{n-1}) dx_1 \dots dx_{n-1} \times \\ \times v_n(t_n, x_n | t_1, x_1; \dots; t_{n-1}, x_{n-1}) dx_n. \end{aligned} \quad (15.1)$$

Множитель  $v_n dx_n$  в правой части есть условная вероятность того, что в момент  $t_n$  значение случайной функции  $\xi(t)$  окажется в интервале  $(x_n, x_n + dx_n)$  *при условии*, что в предшествующие моменты  $t_i$  она принимала значения, равные  $x_i$  ( $i = 1, \dots, n-1$ ). Таким образом, условная вероятность  $v_n$  *состояния*  $x_n$  в момент  $t_n$  зависит от предшествующего пути, пройденного случайной функцией, от множества предшествую-

<sup>1)</sup> В дальнейшем мы будем иметь дело главным образом с *процессами* — случайными функциями непрерывного аргумента. Общая теория случайных процессов марковского типа была развита в начале 30-х годов А. Н. Колмогоровым.

щих состояний, из которых, в зависимости от числа измерений  $n$ , выбраны  $n - 1$  состояний  $x_1, \dots, x_{n-1}$ .

Если обозначить через  $\xi_p(\theta)$  значение  $\xi(t)$ , фактически принятое в момент  $\theta \leq t_0$ , то для всякого  $t > t_0$  распределение, вообще говоря, будет зависеть от всего предшествующего течения процесса  $\xi_p(\theta)$ . Эту ситуацию иллюстрирует рис. 8,

на котором  $v$  построена в функции от  $x$  и  $t$ . Над кривой  $\xi_p(\theta)$ , лежащей на плоскости  $(x, t)$ , — кривой фактически принятых до момента  $t_0$  значений — возвышается бесконечно высокий «забор» дельта-функций. Ход  $v \parallel x$  для любого  $t > t_0$  зависит от всей формы этого «забора», т. е. будучи функцией от  $x$  и  $t$ , является функционалом от  $\xi_p(\theta)$ . Если при том же значении  $x_0$  в момент  $t_0$  взята иной вид  $\xi_p(\theta)$ , т. е. иначе изогнуть «забор», то форма поверхности  $v(x, t)$  над участком  $t > t_0$  плоскости  $(x, t)$  изменится. Говорят, что  $\xi(t)$  испытывает *вероятностное последствие* со стороны ранее принятых значений.

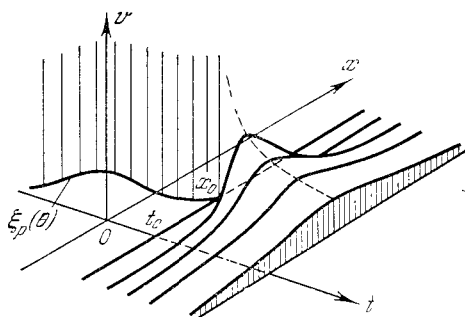


Рис. 8.

Разные участки кривой  $\xi_p(\theta)$  могут влиять на последующее распределение вероятности в различной степени. В частности, процесс может обладать затухающим последствием: значения  $\xi_p(\theta)$  оказывают тем меньшее влияние, чем более давними они являются. Это «забывание» может быть не только постепенным, но и резко наступающим. Например, может отсутствовать влияние всех  $\xi_p(\theta)$  при  $\theta < t_0 - \tau$ , т. е. последствие оказывают только значения, принятые на отрезке времени  $\tau$  до момента  $t_0$ . Еще более частный, но физически важный случай — тот, когда для всех  $t > t_0$  условная вероятность однозначно определяется значением  $x_0$ , принятым в момент  $t_0$ , и совсем не зависит от предшествующей истории, т. е. имеет вид  $v_2(t, x|t_0, x_0)dx$ . Это и есть процесс марковского типа, или процесс без последствия.

Заметим, что такой процесс можно рассматривать как непосредственное обобщение динамической закономерности, описываемой дифференциальным уравнением движения первого порядка. Решение этого уравнения, удовлетворяющее условию  $x = x_0$  при  $t = t_0$ , есть

$$x = f(t, x_0, t_0), \quad \text{причем} \quad x_0 \equiv f(t_0, x_0, t_0).$$

Детерминированная функция  $f$  однозначно определяет значение  $x$  в момент  $t$  при заданном значении  $x_0$  в момент  $t_0$ . Можно сказать поэтому, что плотность условной вероятности значения  $x$  в момент  $t$  выражается в этом частном случае дельта-функцией:

$$v_2(t, x | t_0, x_0) = \delta[x - f(t, x_0, t_0)].$$

Примером такого динамического процесса может служить движение частицы в вязкой жидкости, реакция которой учитывается в виде стокового трения, т. е. силой  $F = -hu$ , где  $u$  — скорость частицы. Мы имеем тогда для  $u$  (это и есть величина  $x$  в данном примере) уравнение движения  $m\dot{u} + hu = 0$ , из которого следует

$$u = u_0 e^{-\frac{h}{m}(t-t_0)} \equiv f(t, u_0, t_0).$$

Известно, однако, что выражение  $F = -hu$  для силы трения представляет собой лишь первое приближение, пригодное только для малых ускорений. В общем случае те возмущения (вязкие волны), которые создаются в жидкости движением частицы, влияют на испытываемую ею реакцию в течение длительного времени. В результате реакция в какой-либо момент  $t$  зависит не просто от скорости  $u$  частицы в этот же момент, а от всего ее *предшествующего* движения (вязкое последствие). Следовательно, при более точной постановке задачи уравнение движения частицы в вязкой жидкости будет уже интегро-дифференциальным и состояние  $u$  в момент  $t$  не будет определяться однозначно состоянием  $u_0$  в момент  $t_0$  (§ 35).

Нетрудно представить себе соответствующий переход к статистическому явлению, например брауновскому блужданию частицы в жидкости. При реакции стокового типа это марковский процесс, а с учетом вязкого последствия появится вероятностное последствие: *распределение вероятности  $u$  в момент  $t$  будет зависеть от всего предшествующего движения частицы.*

Вернемся к общей формуле (15.1). Для процесса марковского типа плотность условной вероятности состояния  $(x_n, t_n)$  относительно цепочки предыдущих состояний  $(x_{n-1}, t_{n-1}), \dots, (x_1, t_1)$  совпадает, согласно сказанному, с плотностью условной вероятности относительно последнего из предыдущих состояний  $(x_{n-1}, t_{n-1})$  и не зависит от более ранней истории. Другими словами, статистические свойства процесса в момент  $t_n$  определены состоянием в момент  $t_{n-1}$  и никак не зависят от любой информации о течении процесса до момента  $t_{n-1}$ :

$$v_n(t_n, x_n | t_{n-1}, x_{n-1}; \dots; t_1, x_1) = v_2(t_n, x_n | t_{n-1}, x_{n-1})$$

для любых  $t_1 < t_2 < \dots < t_n$ ). Связь между  $w_n$  и  $w_{n-1}$  имеет, следовательно, вид

$$\begin{aligned} w_n(t_1, x_1; \dots; t_n, x_n) &= \\ &= w_{n-1}(t_1, x_1; \dots; t_{n-1}, x_{n-1}) v_2(t_n, x_n | t_{n-1}, x_{n-1}). \end{aligned}$$

Применяя эту формулу последовательно к  $w_{n-1}$ ,  $w_{n-2}$  вплоть до  $w_2$ :

$$w_2(t_1, x_1; t_2, x_2) = w_1(t_1, x_1) v_2(t_2, x_2 | t_1, x_1), \quad (15.2)$$

получаем

$$\begin{aligned} w_n(t_1, x_1; \dots; t_n, x_n) &= \\ &= w_1(t_1, x_1) v_2(t_2, x_2 | t_1, x_1) v_2(t_3, x_3 | t_2, x_2) \dots \\ &\dots v_2(t_n, x_n | t_{n-1}, x_{n-1}). \end{aligned} \quad (15.3)$$

Таким образом, для того чтобы написать  $n$ -мерную плотность вероятности марковского процесса, надо знать только *две* функции — одномерную плотность вероятности  $w_1$  и плотность условной вероятности  $v$  (индекс 2 при  $v$  мы далее опустим).

Соотношение (15.3), конечно, тоже может быть положено в основу определения марковского процесса. Если оно имеет место, то принято называть  $v$  *вероятностью перехода* (из предыдущего состояния в последующее), что оправдано именно отсутствием последействия.

Выражение (15.2) для *двумерного* распределения может быть написано для любого процесса — просто по определению условной вероятности. Но уже для  $n = 3$  существование вероятностей перехода означает сильное ограничение вида возможных функций распределения<sup>2)</sup>.

Итак, случайная функция марковского типа полностью определяется заданием одномерной функции распределения и вероятности перехода между любыми двумя  $t_1$  и  $t_2$ , тогда как

<sup>1)</sup> Иногда этот случай называют марковским процессом *первого порядка*, понимая под марковским процессом  $k$ -го порядка процесс, у которого условная вероятность состояния зависит от фиксированного числа  $k$  предшествующих состояний. Например, для процесса второго порядка она имеет вид  $v_3(t_3, x_3 | t_2, x_2; t_1, x_1)$ . Эти более общие случаи мы рассматривать не будем.

<sup>2)</sup> Это ограничение можно привести к некоторому интегральному соотношению для  $w_n$ . Возьмем случай  $n = 3$ . Учитывая, что

$$v(x_2 | x_1) = \frac{w_2(x_1, x_2)}{w_1(x_1)}, \quad v(x_3 | x_2) = \frac{w_2(x_2, x_3)}{w_1(x_2)},$$

и выражая  $w_2$  и  $w_1$  через интегралы по «лишним» переменным от  $w_3(x_1, x_2, x_3)$ , получаем из равенства  $w_3 = w_1(x_1) v(x_2 | x_1) v(x_3 | x_2)$  *условие*

$$w_3 \int w_3 dx_1 dx_3 = \int w_3 dx_3 \int w_3 dx_1.$$

в общем случае необходимо задание всех конечномерных распределений.

Случайная функция *марковского* типа называется *однородной* по  $t$ , если вероятность перехода  $v(t_2, x_2 | t_1, x_1)$  зависит только от  $t_2 - t_1$ . Тогда

$$\omega_2(t_1, x_1; t_2, x_2) = \omega_1(t_1, x_1) v(x_2 | t_2 - t_1, x_1). \quad (15.4)$$

Одномерное распределение, вообще говоря, зависит от  $t_1$ , так что двумерное зависит от  $t_1$  и  $t_2$  в отдельности. Возможны, однако, случаи, когда существует *стационарное* одномерное распределение  $\omega(x)$ , одинаковое для всех  $t$ . Тогда

$$\omega_2(t_1, x_1; t_2, x_2) = \omega(x_1) v(x_2 | t_2 - t_1, x_1) = \omega_2(x_1, t_2 - t_1, x_2), \quad (15.5)$$

т. е. двумерное распределение зависит теперь только от  $t_2 - t_1$ . При этом, поскольку распределение  $\omega$  должно иметь место при любом  $t$ , необходимо

$$\int \omega_2(x_1, t_2 - t_1, x_2) dx_1 = \omega_1(t_2 - t_1, x_2) = \omega(x_2),$$

т. е. распределение  $\omega(x)$  должно удовлетворять условию

$$\int \omega(x_1) v(x_2 | t_2 - t_1, x_1) dx_1 = \omega(x_2).$$

В этом случае однородная случайная функция марковского типа является *стационарной*, в соответствии с определением стационарности, которое мы приведем далее.

## § 16. Стационарные процессы

Представление о стационарном процессе может дать, например, наблюдение каких-либо флуктуаций при неизменных макроскопических условиях. Электрический шум в сопротивлении, температура которого постоянна; турбулентность в установившемся потоке; дробовой эффект при постоянном анодном токе и т. п. — все это стационарные случайные процессы. Практически достаточно наблюдения постоянства условий в течение конечного промежутка времени, к началу которого возможные переходные процессы в рассматриваемой системе уже закончились, а в конце еще не начались.

Все вероятностные характеристики стационарной случайной функции не должны, таким образом, меняться при изменении начала отсчета времени. Это значит, что

$$\omega_n(t_1, x_1; \dots; t_n, x_n) \equiv \omega_n(t_1 + \tau, x_1; \dots; t_n + \tau, x_n),$$

т. е.  $\omega_n$  (а значит, и соответствующие характеристические функции) могут зависеть только от *разностей* значений  $t_1, \dots, t_n$ ,



но не от самих этих значений порознь. Про случайную функцию, удовлетворяющую этому условию, говорят, что она *одно-родна по аргументу  $t$* . Этот общий термин используется обычно в том случае, когда под  $t$  понимается пространственная координата (например, функция, однородная по абсциссе). Если же параметр  $t$  есть время, то однородность по времени большей частью называется *стационарностью*.

Из определения стационарной случайной функции вытекает, что ее одномерная плотность вероятности  $w_1$  вообще не зависит от  $t$ , двумерная плотность  $w_2$  может зависеть только от  $t_2 - t_1$ , трехмерная — от  $t_2 - t_1$  и  $t_3 - t_1$  и т. д. Следовательно, и условная вероятность  $x_2$  относительно  $x_1$  для стационарной функции тоже зависит только от  $t_2 - t_1$ :

$$v(t_2, x_2 | t_1, x_1) = \frac{w_2(x_1, t_2 - t_1, x_2)}{w_1(x_1)} = v(x_2 | t_2 - t_1, x_1).$$

Во многих случаях у условной вероятности существует стационарный предел при возрастании  $\tau = t_2 - t_1$ :

$$\lim_{\tau \rightarrow \infty} v(x_2 | \tau, x_1) = \frac{1}{w_1(x_1)} \lim_{\tau \rightarrow \infty} w_2(x_1, \tau, x_2) = w_1(x_2),$$

т. е. зависимость распределения в момент  $t_1 + \tau$  от значения, принятого в момент  $t_1$ , при  $\tau \rightarrow \infty$  ослабевает и в пределе исчезает. Условная вероятность переходит в стационарную одномерную вероятность состояния. Это обеспечивает так называемую *эргодичность* процесса, очень важное свойство, которое мы рассмотрим в дальнейшем (§ 20).

Подчеркнем еще раз, что марковские и стационарные процессы не представляют собой двух не перекрывающихся классов. Для процесса без последствия  $n$ -мерная плотность вероятности имеет вид (15.3), так что, вообще говоря, стационарности нет. Она будет иметь место, как уже было сказано, тогда, когда, во-первых, марковский процесс однороден (т. е. вероятность перехода  $v$  от  $x_1$  к  $x_2$  зависит только от  $t_2 - t_1$ ) и, во-вторых, существует стационарная одномерная функция распределения  $w_1(x)$ . С другой стороны, стационарный процесс описывается конечномерными плотностями вероятности  $w_n$ , которые зависят только от разностей  $t_i - t_h$ , но которые при  $n > 2$  в общем случае непредставимы через вероятности перехода  $v$  между двумя последовательными состояниями.

Все приведенные выше определения и свойства легко распространяются, по аналогии с многомерными случайными величинами, на *многомерные случайные функции*, т. е. совокупности любого числа  $N$  одномерных случайных функций  $\rho(t) = \{\xi^{(1)}(t), \dots, \xi^{(N)}(t)\}$ . Обозначая возможные значения  $\rho(t)$  в какой-либо момент времени  $t_i$  через  $\mathbf{r}_i$ , мы можем повторить

все сказанное ранее применительно к  $Nn$ -мерному распределению

$$\omega_n(t_1, \mathbf{r}_1; \dots; t_n, \mathbf{r}_n) d\mathbf{r}_1 \dots d\mathbf{r}_n.$$

Для марковского многомерного процесса  $\omega_n$  выражается через вероятность состояния  $\omega_1(t_1, \mathbf{r}_1)$  и вероятности перехода  $v(t_2, \mathbf{r}_2 | t_1, \mathbf{r}_1)$ . Для стационарной функции плотности вероятности  $\omega_n$  зависят только от разностей  $t_i$ ; смотря по тому, меняется ли  $t$  дискретно или непрерывно, можно различать многомерные случайные последовательности и процессы и т. д.

### § 17. Моменты случайных функций

Моменты случайной функции  $\xi(t)$  вводятся совершенно так же и с той же целью, что и для системы случайных величин. Это средние значения вида

$$\overline{\xi(t_1) \dots \xi(t_n)},$$

причем произвольно выбираемые  $t_1, \dots, t_n$  могут и совпадать между собой как частично ( $m$  различных значений  $t_i$ ,  $m \leq n$ ), так и полностью ( $t_1 = \dots = t_n$ , т. е.  $m = 1$ ). Число сомножителей  $n$  называется *порядком* момента, так что  $\overline{\xi(t)}$  есть момент первого порядка,  $\overline{\xi(t_1)\xi(t_2)}$ ,  $\overline{\xi^2(t_1)}$ ,  $\overline{\xi^2(t_2)}$  — моменты второго порядка (смешанный и два средних квадрата) и т. д. Существенное состоит в том, что моменты случайных величин являются просто числами, в то время как моменты случайных функций представляют собой функции выбранных значений параметра  $t$ , поскольку от этих значений зависят функции распределения.

Очевидно, для вычисления момента любого порядка  $n$  достаточно  $m$ -мерного распределения, где  $m$  — по-прежнему число различных значений  $t_i$  ( $1 \leq m \leq n$ ). Например, момент  $n$ -го порядка  $\overline{\xi^n(t_1)}$  определяется при помощи одномерного распределения:

$$\overline{\xi^n(t_1)} = \int x_1^n \omega_1(t_1, x_1) dx_1 = M_n(t_1). \quad (17.1)$$

Смешанный момент того же порядка  $n$ , содержащий  $t_1$  и  $t_2$ , — при помощи двумерного распределения:

$$\begin{aligned} \overline{\xi^\alpha(t_1) \xi^\beta(t_2)} &= \int x_1^\alpha x_2^\beta \omega_2(t_1, x_1; t_2, x_2) dx_1 dx_2 = \\ &= M_{\alpha+\beta}(t_1, t_2), \quad \text{где } \alpha + \beta = n, \end{aligned} \quad (17.2)$$

и т. д. вплоть до

$$\overline{\xi(t_1) \xi(t_2) \dots \xi(t_n)} = \int x_1 x_2 \dots x_n \omega_n dx_1 \dots dx_n = M_n(t_1, \dots, t_n). \quad (17.3)$$

Это вытекает, конечно, из согласованности всех конечномерных распределений: выполняя усреднение  $n$  сомножителей, из которых только  $m \leq n$  различны, мы можем воспользоваться любой функцией распределения  $w_N$  с  $N \geq m$ , но интегрирование по «лишним» переменным тотчас же сведет ее к  $w_m$ . Например,

$$\begin{aligned} \int x_1^\alpha x_2^\beta w_N dx_1 \dots dx_N &= \int x_1^\alpha x_2^\beta dx_1 dx_2 \int w_N dx_3 \dots dx_N = \\ &= \int x_1^\alpha x_2^\beta w_2 dx_1 dx_2. \end{aligned}$$

Полагая в (17.3) значения  $t_i$  в том или ином количестве одинаковыми, мы можем, очевидно, получить из (17.3) полный набор всех моментов  $n$ -го порядка. В дальнейшем мы часто будем пользоваться более лаконичными обозначениями для моментов:

$$\overline{x_1^n}, \overline{x_1^\alpha x_2^\beta}, \overline{x_1 x_2 \dots x_n} \text{ и т. д.}$$

Центральные моменты определяются как средние значения пульсаций (или флуктуаций)  $\tilde{\xi}(t_i) = \xi(t_i) - \bar{\xi}(t_i)$ . Аналогичное (17.3), выражение для центрального момента  $n$ -го порядка есть

$$\begin{aligned} \overline{\tilde{\xi}(t_1) \dots \tilde{\xi}(t_n)} &= \overline{[\xi(t_1) - \bar{\xi}(t_1)] \dots [\xi(t_n) - \bar{\xi}(t_n)]} = \\ &= \overline{(x_1 - \bar{x}_1)(x_2 - \bar{x}_2) \dots (x_n - \bar{x}_n)}, \end{aligned}$$

что нетрудно, конечно, записать в виде алгебраической суммы произведений простых моментов порядка от  $n$  до 1.

Относительно роли моментов случайной функции можно сказать то же, что и о роли моментов случайных величин. Конечно, они дают менее полную характеристику случайной функции, чем ее функции распределения, но, во-первых, в ряде практически важных вопросов эта характеристика оказывается достаточной, а, во-вторых, само нахождение моментов зачастую значительно проще, чем вычисление конечномерных функций распределения. Кроме того, для практических целей, а также в ряде вопросов теории наиболее существенную роль играют моменты низших порядков, в особенности первого и второго.

Момент первого порядка, или *среднее значение* функции, есть

$$\overline{\xi(t)} = \int x w_1(t, x) dx \equiv \bar{\xi}(t). \quad (17.4)$$

Моменты второго порядка, для которых удобно ввести специальные обозначения, — это *смешанный момент*:

$$\overline{\xi(t_1) \xi(t_2)} = \int x_1 x_2 w_2(t_1, x_1; t_2, x_2) dx_1 dx_2 \equiv B(t_1, t_2), \quad (17.5)$$

равный при совпадающих аргументах  $t_1 = t_2 = t$  среднему квадрату функции в момент времени  $t$ :

$$\overline{\xi^2(t)} = B(t, t) \equiv \bar{\xi}^2(t), \quad (17.6)$$

и центральный момент второго порядка, называемый *функцией корреляции* (иногда — автокорреляции)<sup>1)</sup>:

$$\begin{aligned} [\xi(t_1) - \bar{\xi}(t_1)] [\xi(t_2) - \bar{\xi}(t_2)] = \\ = B(t_1, t_2) - \bar{\xi}(t_1) \bar{\xi}(t_2) \equiv \psi(t_1, t_2), \end{aligned} \quad (17.7)$$

равный при  $t_1 = t_2 = t$  дисперсии  $\xi(t)$  в момент времени  $t$ :

$$\bar{\xi}^2(t) - \bar{\xi}^2(t) = \psi(t, t) \equiv D[\xi(t)]. \quad (17.8)$$

Среднее значение определяет «центр тяжести» распределения, но, в отличие от случайной величины, теперь — в каждый момент времени  $t$ . Дисперсия является простейшей мерой распыленности, или разброса, относительно центра тяжести (опять-таки в каждый момент времени  $t$ ). Наконец, функция корреляции служит характеристикой *статистической связи* между значениями случайной функции, принимаемыми ею в какие-либо два различных момента времени  $t_1$  и  $t_2$ . Этой последней цели служит также *коэффициент корреляции*, определяемый как

$$K(t_1, t_2) = \frac{\psi(t_1, t_2)}{\sqrt{\psi(t_1, t_1) \psi(t_2, t_2)}} = \frac{\overline{\xi(t_1) \xi(t_2)} - \bar{\xi}(t_1) \bar{\xi}(t_2)}{\sqrt{D[\xi(t_1)] D[\xi(t_2)]}}, \quad (17.9)$$

т. е. пропорциональный функции корреляции и нормированный к единице при  $t_1 = t_2$ . Если  $\xi(t_1)$  и  $\xi(t_2)$  статистически независимы, т. е. двумерное распределение распадается на произведение двух одномерных:

$$\omega_2(t_1, x_1; t_2, x_2) = \omega_1(t_1, x_1) \omega_1(t_2, x_2),$$

то

$$B(t_1, t_2) = \overline{\xi(t_1) \xi(t_2)} = \bar{\xi}(t_1) \bar{\xi}(t_2),$$

в силу чего функция и коэффициент корреляции обращаются в нуль. *Статистическая независимость влечет за собой отсутствие корреляции*, но, как и для случайных величин, обратное утверждение в общем случае неверно.

<sup>1)</sup> К сожалению, общепринятая терминология до сих пор еще не установилась. В английской литературе  $\psi(t_1, t_2)$  большей частью именуется ковариацией (covariance — в соответствии с тем, что дисперсия называется variance), а функцией корреляции (correlation function) называется момент  $B(t_1, t_2)$  и коэффициент корреляции  $K(t_1, t_2)$  (см. ниже). В последнее время, в связи с развитием корреляционной теории когерентности колебаний и волн, смешанный момент  $B(t_1, t_2)$  часто называют в оптике *функцией когерентности* (второго порядка) и используют для него обозначение  $\Gamma^{(2)}(t_1, t_2)$ .

Заметим, что  $B(t_1, t_2)$  всегда можно при помощи (15.2) представить в виде

$$\begin{aligned} B(t_1, t_2) &= \int x_1 x_2 \omega_2 dx_1 dx_2 = \\ &= \int x_1 \omega_1(t_1, x_1) \bar{dx}_1 \int x_2 v(t_2, x_2 | t_1, x_1) dx_2. \end{aligned}$$

Внутренний интеграл представляет собой *условное среднее* от  $\xi(t_2)$ , т. е. среднее значение в момент времени  $t_2$  при условии, что в момент времени  $t_1$  функция приняла значение  $x_1$ . Обозначив это условное среднее, зависящее от  $t_1, t_2$  и  $x_1$ , через

$$\bar{x}_2(t_2 | t_1, x_1),$$

мы получаем, таким образом,

$$B(t_1, t_2) = \int x_1 \bar{x}_2(t_2 | t_1, x_1) \omega_1(t_1, x_1) dx_1. \quad (17.10)$$

Для многомерной случайной функции  $\rho(t) = \{\xi^{(1)}(t), \dots, \xi^{(N)}(t)\}$  в рассмотрение входят моменты, содержащие произведения разных компонент  $\xi^{(h)}(t)$ . Если не идти дальше моментов второго порядка, то для многомерной функции, наряду со средними значениями  $\bar{\xi}^{(h)}(t)$ , приходится иметь дело не только с моментами вида  $B_{hk}(t_1, t_2) = \overline{\xi^{(k)}(t_1) \xi^{(h)}(t_2)}$  и соответственно с функциями автокорреляции  $\psi_{hk}(t_1, t_2)$ , но и со *взаимными* моментами второго порядка, содержащими произведения двух различных функций  $\xi^{(h)}(t)$ . Мы получаем, таким образом, *матрицу* моментов второго порядка с элементами

$$B_{ik}(t_1, t_2) = \overline{\xi^{(i)}(t_1) \xi^{(k)}(t_2)}, \quad i, k = 1, 2, \dots, N, \quad (17.11)$$

и *корреляционную матрицу* с элементами

$$\begin{aligned} \psi_{ik}(t_1, t_2) &= [\bar{\xi}^{(i)}(t_1) - \bar{\xi}^{(i)}(t_1)] [\bar{\xi}^{(k)}(t_2) - \bar{\xi}^{(k)}(t_2)] = \\ &= B_{ik}(t_1, t_2) - \bar{\xi}^{(i)}(t_1) \bar{\xi}^{(k)}(t_2). \end{aligned} \quad (17.12)$$

## § 18. Корреляционная теория

Теория случайных функций, рассматривающая только одномерное и двумерное распределения, называется *корреляционной теорией* таких функций. Распределения высшего числа измерений она оставляет вне поля зрения. Соответственно в пределах корреляционной теории мы имеем возможность находить моменты различных порядков, но содержащие не более двух значений параметра  $t$ , в том числе среднее значение случайной функции и ее функцию корреляции. Вообще, мы можем вычислить среднее значение любой детерминированной функции вида

$F[\xi(t_1), \xi(t_2)]$ , если, конечно, это среднее существует:

$$\overline{F[\xi(t_1), \xi(t_2)]} = \int F(x_1, x_2) w_2(t_1, x_1; t_2, x_2) dx_1 dx_2.$$

Что можно сказать о моментах первого и второго порядков специально для марковских и для стационарных процессов?

Очевидно, марковость процесса не вносит здесь ничего специфического, так как двумерное распределение *всегда* может быть выражено по (15.2) через условную вероятность  $v(t_2, x_2 | t_1, x_1)$ . То, что для марковского процесса  $v$  является вероятностью перехода, существенно лишь для  $n$ -мерных распределений с  $n \geq 3$ , но этих распределений корреляционная теория не рассматривает.

Напротив, стационарность процесса существенно сказывается на интересующих нас моментах первого и второго порядков. Действительно, поскольку одномерная плотность вероятности  $w_1(x)$  не зависит в этом случае от  $t$ , моменты любого порядка, вычисленные для *одного* значения  $t$ , тоже не будут зависеть от  $t$ , т. е. будут числовыми константами:

$$\overline{\xi^n(t)} = \int x^n w_1(x) dx = \overline{\xi^n} = \text{const.}$$

В частности, постоянными будут среднее значение  $\overline{\xi}$ , средний квадрат  $\overline{\xi^2}$  и дисперсия  $D[\xi] = \overline{\xi^2} - \overline{\xi}^2$ . Далее, поскольку двумерная плотность вероятности стационарного процесса зависит от  $t_1$  и  $t_2$  только через разность  $\tau = t_2 - t_1$ , смешанный момент второго порядка, а значит, и функция корреляции тоже будут зависеть только от  $\tau$ :

$$\begin{aligned} \overline{\xi(t_1) \xi(t_2)} &= B(\tau), \quad \overline{\xi^2} = B(0), \\ B(\tau) - \overline{\xi}^2 &= \psi(\tau), \quad D[\xi] = \psi(0), \\ K(\tau) &= \frac{\psi(\tau)}{\psi(0)} = \frac{\psi(\tau)}{D[\xi]}. \end{aligned} \quad (18.1)$$

Заметим, что  $B(\tau)$  [а тем самым  $\psi(\tau)$  и  $K(\tau)$ ] — обязательно четная функция  $\tau$ . Это следует из инвариантности по отношению к сдвигу начала отсчета времени  $t$  на  $\tau$ :

$$B(\tau) = \overline{\xi(t) \xi(t + \tau)} = \overline{\xi(t - \tau) \xi(t)} = B(-\tau).$$

Так как постоянное среднее значение обычно является легко измеряемой величиной, во многих случаях можно рассматривать только флуктуацию  $\tilde{\xi} = \xi - \overline{\xi}$ , т. е., не теряя общности, предполагать, что  $\overline{\xi} = 0$ . Тогда  $\psi(\tau)$  совпадает с  $B(\tau)$ :

$$\psi(\tau) = B(\tau) = \overline{\xi(t) \xi(t + \tau)}, \quad (18.2)$$

а коэффициент корреляции будет

$$K(\tau) = \frac{\psi(\tau)}{\psi(0)} = \frac{\xi(t)\xi(t+\tau)}{\xi^2}.$$

Для стационарных случайных функций особая роль функции корреляции связана еще и с тем, что она определяет *гармонический спектр* процесса, но к этому мы обратимся позднее.

В пределах корреляционной теории возможно более широкое определение стационарности. *Стационарными в широком смысле* (просто стационарными по А. Я. Хинчину [1]) называются случайные функции  $\xi(t)$ , у которых *среднее значение постоянно*:  $\xi(t) = \bar{\xi} = \text{const}$ , а *момент второго порядка В зависит только от*  $\tau = t_2 - t_1$  *и конечен при*  $\tau = 0$ , т. е. конечен средний квадрат  $\bar{\xi}^2 = B(0)$ . Тем самым конечно и постоянное среднее значение  $\bar{\xi}$ , поскольку  $D[\xi] = \bar{\xi}^2 - \bar{\xi}^2 \geq 0$ . Прежнее определение, предполагавшее инвариантность всех конечномерных распределений при сдвиге начала отсчета  $t$  (*стационарность в узком смысле*), не содержало требования конечности моментов первых двух порядков и в этом смысле было менее жестким. Но стационарность в широком смысле не означает, что все  $\omega_n$  при  $n > 2$  инвариантны при изменении начала отсчета  $t$ , и в этом отношении она действительно шире.

Посмотрим, при каких условиях может быть стационарной в широком смысле функция  $\xi(t) = A \cos(\omega t + \theta)$ , где в общем случае и  $A$ , и  $\theta$  — случайные величины.

Если случайна только амплитуда  $A$ , а фаза  $\theta$  имеет какое-то фиксированное значение, то

$$\overline{\xi(t)} = \bar{A} \cos(\omega t + \theta)$$

и условие постоянства среднего значения выполняется лишь при  $\bar{A} = 0$ , когда и  $\overline{\xi(t)} = 0$ . Для смешанного момента получаем

$$\begin{aligned} \overline{\xi(t)\xi(t+\tau)} &= \overline{A^2 \cos(\omega t + \theta) \cos[\omega(t+\tau) + \theta]} = \\ &= \frac{\bar{A}^2}{2} [\cos(2\omega t + \omega\tau + 2\theta) + \cos \omega\tau], \end{aligned}$$

т. е.  $B = B(t, \tau)$ , и, следовательно, даже при  $\bar{A} = 0$  стационарности нет.

Пусть теперь фаза  $\theta$  тоже случайна, причем независима от  $A$ . Тогда

$$\overline{\xi(t)} = \bar{A} \overline{\cos(\omega t + \theta)} = \bar{A} (\cos \omega t \overline{\cos \theta} - \sin \omega t \overline{\sin \theta}).$$

Очевидно, независимость  $\overline{\xi(t)}$  от  $t$ , т. е. равенство  $\overline{\xi(t)} = 0$ , можно обеспечить теперь не только при  $\bar{A} = 0$ , но и при

$\overline{\cos \theta} = \overline{\sin \theta} = 0$ , что будет иметь место, если плотность вероятности фазы  $w_\theta(\theta)$  ортогональна в интервале  $(0, 2\pi)$  к  $\cos \theta$  и  $\sin \theta$ , т. е. представима рядом Фурье:

$$w_\theta(\theta) = \frac{1}{2\pi} + \sum_{n=2}^{\infty} (a_n \cos n\theta + b_n \sin n\theta) \geq 0.$$

Смешанный момент

$$\overline{\xi(t) \xi(t+\tau)} = \frac{\overline{A^2}}{2} [\overline{\cos(2\omega t + \omega\tau + 2\theta)} + \cos \omega\tau]$$

теперь тоже может быть сделан функцией только от  $\tau$ . Надо только потребовать  $\overline{\cos 2\theta} = \overline{\sin 2\theta} = 0$ , т. е. чтобы в ряде Фурье, выражающем  $w_\theta(\theta)$ , отсутствовали также члены с  $n = 2$ . Тогда

$$\overline{\xi(t) \xi(t+\tau)} = \frac{\overline{A^2}}{2} \cos \omega\tau = B(\tau).$$

Таким образом, при случайных и независимых  $A$  и  $\theta$  стационарность случайной функции  $\xi(t) = A \cos(\omega t + \theta)$  в широком смысле будет обеспечена, если  $\overline{A^2}$  конечно, а распределение  $\theta$  имеет вид

$$w_\theta(\theta) = \frac{1}{2\pi} + \sum_{n=3}^{\infty} (a_n \cos n\theta + b_n \sin n\theta).$$

Нетрудно понять, что при наличии в этом ряде Фурье только гармоник с  $n > N$  все смешанные моменты  $\xi(t)$  до порядка  $N$  включительно будут зависеть только от разностей соответствующих моментов времени. В частности, если распределение равномерно в интервале  $(0, 2\pi)$ :

$$w_\theta(\theta) = 1/2\pi,$$

то сказанное относится к моментам любого порядка, что равносильно стационарности  $\xi(t)$  в узком смысле.

Как сказано, оперирование только распределениями  $w_1$  и  $w_2$  часто оказывается практически достаточным, хотя оно и не заменяет в общем случае полного задания случайной функции. Существует, однако, очень важное исключение, благодаря которому ценность корреляционной теории существенно возрастает.

Довольно часто рассматриваемые случайные функции принадлежат к классу *нормальных* (или гауссовых) процессов, т. е. все  $w_n$  представляют собой  $n$ -мерные гауссовы распределения. Это имеет место в силу того, что выполнены условия централь-



ной предельной теоремы, как, например, в явлении дробового шума (§ 10), в распределении скоростей микрочастиц [максвелловское распределение (§ 7)] и т. д. Но  $n$ -мерное гауссово распределение содержит в качестве параметров только средние значения и моменты второго порядка (причем и те, и другие всегда конечны), а для нахождения этих моментов достаточно знать лишь  $\omega_2$  (и тем самым  $\omega_1$ ). Таким образом, нормальная случайная функция  $\xi(t)$  вполне определена заданием плотности вероятности  $\omega_2$ . Зная  $\omega_2$ , можно написать любую  $n$ -мерную функцию распределения и, следовательно, ответить на любой вопрос, касающийся какой-либо детерминированной функции от величин  $\xi(t_1), \dots, \xi(t_n)$ .

Ясно также, что для нормальной случайной функции стационарность в широком смысле совпадает со стационарностью в узком смысле, так как из первой следует вторая: если  $\omega_1$  не зависит от  $t$ , а  $\omega_2$  зависит только от  $\tau = t_2 - t_1$ , то все  $\omega_n$  будут инвариантны по отношению к сдвигу начала отсчета  $t$ . Наконец, для нормальных функций некоррелированность означает статистическую независимость, чего нет в общем случае.

Не представляет затруднений распространить любое из определений стационарности на многомерные случайные функции. В корреляционной теории целесообразно по-прежнему понимать стационарность в широком смысле, т. е. подразумевать под стационарностью  $N$ -мерной функции  $\rho(t) = \{\xi^{(1)}(t), \dots, \xi^{(N)}(t)\}$  конечность и постоянство всех  $\overline{\xi^{(h)}(t)}$  и  $[\xi^{(h)}(t)]^2$  и зависимость элементов матрицы  $B_{ih}$  (или корреляционной матрицы  $\psi_{ih}$ ) только от  $\tau = t_2 - t_1$ .

## § 19. Вероятностная сходимость

В дальнейшем нам придется широко оперировать производными и интегралами от случайных процессов. Обе операции — дифференцирование и интегрирование — предполагают, как известно, сходимость некоторой последовательности величин к пределу. Но для случайных величин, задаваемых не детерминированно, а своими распределениями вероятностей, понятие сходимости к пределу (а тем самым и понятия непрерывности, дифференцируемости, интегрируемости для случайных функций) не может обладать тем же смыслом, какой вкладывается в него в анализе. Для последовательности случайных величин  $\xi_N$  ( $N = 1, 2, \dots$ ) возможно лишь *вероятностное* определение сходимости к пределу, что, кстати сказать, открывает и более разнообразные возможности в выборе самого определения. Вероятностная сходимость существенна также и для рассмотрения так называемых *эргодических* свойств случайных функций, к чему мы обратимся в следующем параграфе.

Начнем, для простоты, с рассмотрения различных типов сходимости последовательности случайных величин  $\xi_N$  к (неслучайному) числу  $a$ .

Один из видов вероятностной сходимости — *сходимость в среднем квадратичном* (ср. кв.), под которой понимается обращение в нуль среднего квадратичного отклонения  $\xi_N$  от числа  $a$  при  $N \rightarrow \infty$ :

$$\lim_{N \rightarrow \infty} \langle (\xi_N - a)^2 \rangle = 0, \quad (19.1)$$

что записывают в виде

$$\lim_{N \rightarrow \infty} \xi_N = a \text{ (ср. кв.)}, \quad \text{или} \quad \xi_N \xrightarrow{\text{ср. кв.}} a \text{ при } N \rightarrow \infty,$$

$$\text{или} \quad \text{l.i.m.}_{N \rightarrow \infty} \xi_N = a.$$

Обозначение l. i. m. составлено из начальных букв английского названия этого предела (limit in the mean square). Использование этого вида сходимости наиболее целесообразно в тех случаях, когда приходится иметь дело с квадратичными (в частности, имеющими энергетический смысл) комбинациями случайных величин.

Равенство (19.1) предполагает, очевидно, конечность  $\bar{\xi}_N^2$ . Тем самым конечно и среднее значение  $\bar{\xi}_N$ , поскольку  $D[\xi_N] = \bar{\xi}_N^2 - \bar{\xi}_N^2 \geq 0$ . Вычитая и прибавляя  $\bar{\xi}_N$  в скобках в (19.1), перепишем это равенство иначе:

$$\lim_{N \rightarrow \infty} \langle (\xi_N - \bar{\xi}_N + \bar{\xi}_N - a)^2 \rangle = \lim_{N \rightarrow \infty} \{D[\xi_N] + (\bar{\xi}_N - a)^2\} = 0.$$

Но предел суммы двух неотрицательных величин может быть равен нулю, только если равны нулю пределы обоих слагаемых, т. е.

$$\lim_{N \rightarrow \infty} \bar{\xi}_N = a, \quad \lim_{N \rightarrow \infty} D[\xi_N] = 0.$$

Таким образом,  $a$  — это предел последовательности средних значений  $\bar{\xi}_N$ , а предел дисперсии  $\xi_N$  равен нулю.

Другой вид вероятностной сходимости  $\xi_N$  к  $a$  — *сходимость по вероятности* (по вер.) — определен следующим образом:

$$\lim_{N \rightarrow \infty} P\{|\xi_N - a| < \epsilon\} = 1, \quad (19.2)$$

где, как обычно,  $\epsilon$  — любое сколь угодно малое положительное число. В этом случае пишут

$$\lim_{N \rightarrow \infty} \xi_N = a \text{ (по вер.)} \quad \text{или} \quad \xi_N \xrightarrow{\text{по вер.}} a \text{ при } N \rightarrow \infty.$$

Равенство (19.2) означает, что вероятность попадания  $\zeta_N$  куда-либо вне сколь угодно узкого интервала  $(a - \varepsilon, a + \varepsilon)$  в пределе ( $N \rightarrow \infty$ ) обращается в нуль. Ввиду произвольной малости  $\varepsilon$  это в свою очередь означает, что плотность вероятности  $w_N(x)$  случайной величины  $\zeta_N$  переходит при  $N \rightarrow \infty$  в  $\delta(x - a)$ . Однако отсюда отнюдь не следует, что  $a$  есть предел последовательности  $\bar{\zeta}_N$  и что  $D[\zeta_N]$  стремится к нулю. Более того,  $\bar{\zeta}_N$  и  $\bar{\zeta}_N^2$  могут неограниченно нарастать с увеличением  $N$  или даже быть бесконечными при всяком  $N$ . Пусть, например,  $\zeta_N$  неотрицательны и распределены по закону Коши:

$$w_N(x) = \frac{2N}{\pi(1 + N^2x^2)} \quad (x \geq 0, \quad N = 1, 2, \dots).$$

При всяком  $x > 0$  предел  $w_N(x)$  при  $N \rightarrow \infty$  равен нулю, тогда как  $w_N(0) = 2N/\pi$  и предела не существует. Вместе с тем условие нормировки выполнено всегда:

$$\int_0^\infty w_N(x) dx = \frac{2}{\pi} \operatorname{arctg}(Nx) \Big|_0^\infty = 1,$$

так что  $w_N(x)$  стремится при  $N \rightarrow \infty$  к  $\delta(x)$ . Нетрудно, однако, убедиться, что при *любом*  $N$   $\bar{\zeta}_N$  и  $\bar{\zeta}_N^2$  бесконечны.

Сходимость по вероятности часто называют *сходимостью в смысле закона больших чисел*. Про случайные величины  $\zeta_N$  говорят, что они являются предельно постоянными, если существует такая последовательность постоянных  $a_N$ , что

$$\lim_{N \rightarrow \infty} P\{|\zeta_N - a_N| < \varepsilon\} = 1.$$

Если все  $a_N$  одинаковы (равны  $a$ ), то это равенство переходит в (19.2), т. е. означает, что  $\zeta_N$  сходится по вероятности к  $a$  или же разность  $\zeta_N - a$  сходится по вероятности к нулю.

Сходимость по вероятности следует четко отличать от обычной сходимости

$$\lim_{N \rightarrow \infty} |\zeta_N - a| = 0 \quad \text{или} \quad \zeta_N \rightarrow a.$$

Действительно, относительно поведения *эмпирических чисел* — значений  $\zeta_N$  — математически доказать ничего нельзя. Доказаны могут быть только утверждения, относящиеся к *теоретическим понятиям*, в том числе к понятию вероятности, как оно определено в исходных аксиомах. В сходимости по вероятности речь идет не о том, что  $\zeta_N \rightarrow a$  при  $N \rightarrow \infty$ , а о том, что *вероятность события*  $|\zeta_N - a| < \varepsilon$  стремится к единице. Связь этого утверждения с опытом заключена в «аксиоме измерения», согласно которой вероятность *измеряется* относительной частотой

наступления рассматриваемого случайного события в достаточно длинной серии испытаний, в достаточно обширном ансамбле систем и т. п.

Для лучшего уяснения этой принципиальной стороны вопроса остановимся на некоторых предельных теоремах теории вероятностей, объединяемых под общим названием *закона больших чисел*, а именно на теоремах, относящихся к тому случаю, когда  $\xi_N$  в (19.2) есть среднеарифметическое  $N$  случайных величин  $\xi_k$ :

$$\xi_N = \frac{1}{N} \sum_{k=1}^N \xi_k. \quad (19.3)$$

Мы производим *серию* из  $N$  испытаний, берем их результаты  $\xi_1, \xi_2, \dots, \xi_N$  и вычисляем среднее (19.3). Затем мы смотрим, имеет ли место событие (назовем его событием  $B_N$ ), состоящее в том, что

$$|\xi_N - a| < \varepsilon.$$

Для того чтобы *измерить* вероятность события  $B_N$ , мы должны осуществить *очень большое число  $M$  серий по  $N$  испытаний*, должны иметь *коллектив таких серий*. Закон больших чисел (19.2) утверждает, что чем длиннее серии, образующие коллектив (чем больше  $N$ ), тем ближе  $P\{B_N\}$  к единице, т. е., по «аксиоме измерения», тем большее количество серий будет отвечать наступлению  $B_N$  (в пределе — практически все):

$$\frac{m(B_N)}{M} \rightarrow 1 \quad \text{при } N \rightarrow \infty.$$

Таким образом, это вполне содержательное утверждение, но оно становится таким только при четком сопоставлении математического понятия вероятности с эмпирическим понятием относительной частоты. Без этого закон больших чисел остается некоторой теоремой, логически вытекающей из определенной системы аксиом для величины  $P$ , которая определена как вполне аддитивная, неотрицательная и нормированная к единице функция области.

Зачастую этот вопрос, который мы уже затрагивали в § 1, излагается в учебной литературе довольно сбивчиво, без четкого указания на то, что «аксиома измерения», связывающая понятия теории вероятностей с реальными явлениями, с экспериментом и практикой, *не содержится* в математической теории как таковой. Можно встретить утверждения о том, что фундамент успехов применения теории вероятностей в различных проблемах естествознания и техники заложен именно в законе больших чисел. Если бы это было так, то это означало бы, что

фундамент *практических* успехов есть *логическое следствие* определенных *абстрактных* аксиом и что эти математические аксиомы сами по себе предписывают, как должны вести себя эмпирические величины.

В принципе можно было бы исходить из других аксиом и построить другую теорию вероятностей, выводы которой, будучи иными, чем в существующей теории, были бы столь же логически безупречны и столь же необязательны для реальных явлений. Положение здесь такое же, как и с различными возможными геометриями. Но как только математическая теория *дополняется* определенными способами измерения тех величин, с которыми она оперирует, и становится тем самым физической теорией, ситуация меняется. Правильность или неправильность теории перестает тогда быть вопросом только ее логической непротиворечивости, а становится вопросом ее соответствия реальным вещам и явлениям. Приобретает содержание вопрос об *истинности* самих аксиом, так как теперь это может быть подвергнуто экспериментальной и вообще практической проверке.

Однако еще до такой проверки необходимо внутреннее соответствие между обеими частями физической теории: устанавливаемые способы измерения величин не должны находиться в противоречии с теми уравнениями, которым подчиняют эти величины математическая часть теории. Например, уравнения движения Ньютона предполагают, что сила есть вектор, и поэтому несовместимы с таким способом измерения силы, который характеризовал бы ее только по абсолютной величине. Может быть, в действительности сила не вектор, а скажем, тензор, но это уже *другой* вопрос, касающийся того, насколько хорошо отражает объективную реальность данная физическая теория в целом. Мы же говорим сейчас лишь о том, что наличие противоречия между математической и измерительной частями физической теории делает ее несостоятельной еще до всякой проверки ее следствий на опыте.

С этой точки зрения закон больших чисел отличается от других — логически равносильных ему — теорем теории вероятностей лишь тем, что он, как будет видно из дальнейшего, особенно отчетливо и явно показывает *совместимость* математического определения вероятности и частотного способа ее измерения. Он показывает, что частотная «аксиома измерения» не противоречит математической теории, но последняя, разумеется, не заменяет и не может заменить эту «аксиому».

Доказательство различных теорем, имеющих форму закона больших чисел, использует обычно неравенство Чебышева, доказанное в его диссертации в 1846 г. Пусть случайная величина  $\xi$  имеет конечную дисперсию  $D[\xi]$ . Неравенство Чебышева

утверждает, что

$$P\{|\xi - a| \geq \varepsilon\} \leq \frac{(\overline{(\xi - a)^2})}{\varepsilon^2}. \quad (19.4)$$

Если, в частности,  $a = \bar{\xi}$ , то неравенство (19.4) принимает вид <sup>1)</sup>

$$P\{|\xi - \bar{\xi}| \geq \varepsilon\} \leq \frac{D[\xi]}{\varepsilon^2}. \quad (19.5)$$

Хотя неравенства (19.4) и (19.5) дают лишь весьма грубую оценку  $P$  (более точную оценку можно получить, если известен закон распределения  $\xi$ ), для теоретических построений они очень полезны и важны.

В случае, когда  $\xi$  в неравенстве Чебышева есть среднее арифметическое (19.3) из  $N$  случайных величин  $\xi_i$ , неравенство (19.5) позволяет доказать *теорему Чебышева*, являющуюся довольно общим выражением закона больших чисел. А именно, если  $\xi_i$  — последовательность попарно независимых случайных величин, имеющих равномерно ограниченные дисперсии ( $D[\xi_i] \leq C$ ), то

$$\lim_{N \rightarrow \infty} P\{|\xi_N - \bar{\xi}_N| < \varepsilon\} = 1. \quad (19.6)$$

Действительно,

$$D[\xi_N] = D\left[\frac{1}{N} \sum_1^N \xi_i\right] = \frac{1}{N^2} \sum_1^N D[\xi_i] \leq \frac{C}{N}.$$

Согласно неравенству Чебышева

$$P\{|\xi_N - \bar{\xi}_N| \geq \varepsilon\} \leq \frac{C}{N\varepsilon^2},$$

откуда для вероятности противоположного события и следует теорема (19.6), т. е. сходимость по вероятности  $\xi_N$  к  $\bar{\xi}_N$ .

Частный случай теоремы Чебышева — теорема Пуассона. Пусть  $\xi_i = e_i$  — случайные величины-фиксаторы исхода  $i$ -го испытания [ $e_i = 1$  или  $0$  в соответствии с наступлением или ненаступлением события  $A$  при  $i$ -м испытании, при котором  $P\{A\} = p_i$ ]. Тогда

$$\xi_N = \frac{1}{N} \sum_1^N e_i = \frac{n}{N}, \quad \bar{e}_i = p_i, \quad D[e_i] = p_i q_i \leq \frac{1}{4} \equiv C.$$

---

<sup>1)</sup> Самая общая форма неравенства Чебышева есть  $P\{f(\xi) \geq K\} \leq \frac{\overline{f(\xi)}}{K}$ , где  $f(\xi)$  — неотрицательная детерминированная функция случайной величины  $\xi$ , а  $K$  — произвольное положительное число. Неравенство (19.4) получается отсюда при  $f(\xi) = (\xi - a)^2$  и  $K = \varepsilon^2$ .

Значит,

$$\bar{\xi}_N = \frac{1}{N} \sum_1^N p_i$$

и теорема Чебышева дает

$$\lim_{N \rightarrow \infty} P \left\{ \left| \frac{n}{N} - \frac{1}{N} \sum_1^N p_i \right| < \varepsilon \right\} = 1. \quad (19.7)$$

Это и есть теорема Пуассона. Еще более частный случай — когда  $p_i = p$ . Тогда мы приходим к теореме Бернулли, одной из первых формулировок закона больших чисел:

$$\lim_{N \rightarrow \infty} P \left\{ \left| \frac{n}{N} - p \right| < \varepsilon \right\} = 1. \quad (19.8)$$

Остановимся на этой простейшей форме закона. Теорема (19.8) показывает, что с ростом числа испытаний  $N$  относительная частота события  $A$ , т. е. эмпирическая величина  $n(A)/N$ , сходится *по вероятности* к  $p$  — вероятности события  $A$ . Если бы это было не так, то было бы бессмысленно измерять вероятность при помощи относительной частоты. Но коль скоро это так, то частотный способ измерения вероятностей как  $p$  (по относительной частоте  $n(A)/N$  наступления события  $A$  в серии из  $N$  испытаний), так и  $P$  (по относительной частоте  $m(B_N)/M$  наступления события  $B_N = \left\{ \left| \frac{n}{N} - p \right| < \varepsilon \right\}$  в коллективе из  $M$  серий по  $N$  испытаний) может быть принят в качестве дополнения к математической теории, поскольку он ей не противоречит. После этого уже можно и спрашивать, и проверять на опыте, отражает ли получившаяся в результате физическая теория реальные статистические закономерности.

Любопытно, что для выполнения теоремы (19.8) при всяких значениях  $p$ , т. е. для сходимости по вероятности

$$\lim_{N \rightarrow \infty} \frac{n}{N} = p \text{ (по вер.)},$$

достаточно потребовать, чтобы эта сходимость имела место лишь для  $p \ll 1$  (относительная частота маловероятных событий должна быть мала).

Запишем теперь теорему Чебышева для случая, когда все  $\bar{\xi}_i = a$ . Тогда

$$\bar{\xi}_N = \frac{1}{N} \sum_1^N \bar{\xi}_i = a$$

и теорема принимает вид

$$\lim_{N \rightarrow \infty} P \left\{ \left| \frac{1}{N} \sum_1^N \xi_i - a \right| < \varepsilon \right\} = 1, \text{ т. е. } \lim_{N \rightarrow \infty} \frac{1}{N} \sum_1^N \xi_i = a \text{ (по вер.)},$$

что является основой правила среднего арифметического при измерениях. Отдельные  $\xi_i$  могут сильно отклоняться от  $a$ , но с вероятностью  $P \rightarrow 1$  имеем  $\xi_N \rightarrow a$  при  $N \rightarrow \infty$ . Это происходит потому, что при вычислении среднего значения случайные отклонения отдельных слагаемых компенсируются и в подавляющем большинстве случаев отклонение  $|\xi_N - a|$  оказывается очень малым.

Отклонения  $\xi_i$  от  $a$  могут быть *случайными* ошибками измерения. Но если сама *точность отсчета* при измерении не меньше  $\delta$ , т. е. присутствует *систематическая* ошибка, связанная с ценой деления шкалы, то и точность  $\xi_N$  не меньше  $\delta$  при любом  $N$ , так что бессмысленно, апеллируя к закону больших чисел, стремиться получить и в этом случае значение  $a$  с погрешностью, меньшей  $\delta$ , за счет  $N \rightarrow \infty$ . Довольно широко распространено заблуждение, будто бы среднее арифметическое позволяет превзойти ограниченную снизу точность измерения и получать, скажем, с помощью щиткового амперметра отсчет силы тока с точностью до микроамперов.

Возможна и другая ситуация: сама измеряемая величина может быть случайной (шумовой ток и т. п.). Тогда мы можем быть уверены ( $P \rightarrow 1$ ), что  $\xi_N \rightarrow a$  при  $N \rightarrow \infty$ , т. е. среднее арифметическое стремится к математическому ожиданию случайной величины.

Условие взаимной независимости результатов  $\xi_i$  измерения случайной величины требует, вообще говоря, выполнения ее замеров через достаточно большие промежутки времени. Однако для справедливости закона больших чисел само это условие независимости не необходимо, так как неравенство Чебышева требует лишь  $D[\xi_N] \rightarrow 0$  при  $N \rightarrow \infty$ . Мы не будем останавливаться на более общих теоремах и на необходимых и достаточных условиях, при которых для среднего арифметического  $\xi_N$  (19.3) справедлив закон больших чисел, так как эти условия касаются самой величины  $\xi_N$  и поэтому менее интересны практически, чем более узкие условия, но относящиеся к отдельным слагаемым  $\xi_i$ .

В 1909 г. Э. Борелем (затем — в более общей форме — Ф. П. Кантелли, потом А. Н. Колмогоровым) было доказано более сильное утверждение, чем закон больших чисел. По теореме Бернулли

$$\lim_{N \rightarrow \infty} P \left\{ \left| \frac{n}{N} - p \right| < \varepsilon \right\} = 1.$$



По Борелю (*усиленный закон больших чисел*)

$$P\left\{\lim_{N \rightarrow \infty} \left| \frac{n}{N} - p \right| = 0\right\} = 1, \quad (19.9)$$

т. е. с достоверностью, или, как принято говорить, «почти наверное», относительная частота имеет своим пределом вероятность  $p$ . Это еще более твердое основание для того, чтобы измерять вероятность относительной частотой.

Опираясь на (19.9), можно ввести еще один вид вероятностной сходимости — сходимость в смысле *усиленного закона больших чисел*, которую называют также сходимостью с вероятностью  $P = 1$  или *сходимостью почти наверное* (п. н.):

$$P\left\{\lim_{N \rightarrow \infty} \xi_N = a\right\} = 1. \quad (19.10)$$

Коротко это можно записать в виде

$$\lim_{N \rightarrow \infty} \xi_N = a \text{ (п. н.) или } \xi_N \xrightarrow{\text{п. н.}} a \text{ при } N \rightarrow \infty.$$

Иногда в связи с определением (19.10) возникает недоумение по поводу того, что в нем фигурирует *обычный* предел последовательности *случайных* величин. Создается впечатление, что мы как будто отступаем здесь от высказанного выше утверждения, что сходимость случайных величин может иметь только вероятностный смысл. Но именно об этом идет речь и в данном случае. Среди различных реализаций последовательности  $\xi_N$  возможны и такие реализации, которые сходятся к  $a$  в обычном смысле. Можно показать, что множество таких реализаций обладает определенной вероятностью  $P$  [2]. Сходимость почти наверное означает, что эта вероятность, т. е. вероятность *случайного события*  $\lim_{N \rightarrow \infty} \xi_N = a$ , равна единице. Иначе говоря,

реализации  $\xi_N$ , сходящиеся к  $a$  в обычном смысле, «почти исчерпывают» множество всех возможных реализаций последовательности  $\xi_N$ . Таким образом, мы никуда не уходим в (19.10) от вероятностного определения сходимости, хотя теперь имеется в виду не предел вероятности (как в сходимости по вероятности), а вероятность предела.

Приведем два из условий сходимости  $\xi_N$  к  $a$  почти наверное. Одно из них — *необходимое и достаточное*:

$$\lim_{M \rightarrow \infty} P\{|\xi_N - a| > \varepsilon \text{ хотя бы для одного } N \geq M\} = 0.$$

Однако на практике это условие никогда нельзя проверить. Другое — более сильное *достаточное* условие — состоит в том,

что при каком-либо  $h > 0$  должен сходиться ряд

$$\sum_{N=1}^{\infty} \overline{(\xi_N - a)^h}.$$

Другие достаточные условия и вообще детальную математическую дискуссию вопросов, касающихся вероятностной сходимости, можно найти в книгах [2] (гл. 3) и [3] (гл. 1).

Сходимость в среднем квадратичном влечет за собой (в силу неравенства Чебышева) сходимость по вероятности, а если все  $\xi_N$  почти наверное равномерно ограничены по модулю, то, и обратно, из сходимости по вероятности следует сходимость в среднем квадратичном. Сходимость почти наверное также влечет за собой сходимость по вероятности, но не сходимость в среднем квадратичном; в то же время и сходимость в среднем квадратичном не влечет за собой сходимости почти наверное.

Все сказанное выше о сходимости последовательности случайных величин  $\xi_N$  к числу  $a$  естественным образом распространяется и на сходимость к случайной величине  $\xi$ , определенной на том же множестве возможных значений, что и величины  $\xi_N$ . Так, например, определение сходимости  $\xi_N$  к  $\xi$  в среднем квадратичном есть

$$\lim_{N \rightarrow \infty} \langle (\xi_N - \xi)^2 \rangle = 0, \quad (19.11)$$

а сходимость по вероятности —

$$\lim_{N \rightarrow \infty} P \{ |\xi_N - \xi| < \varepsilon \} = 1. \quad (19.12)$$

Отличие от (19.1) и (19.2) заключается в том, что усреднение в (19.11) и подсчет вероятности  $P$  в (19.12) требуют теперь знания *совместного* распределения величин  $\xi_N$  и  $\xi$ . Если плотность вероятности этого распределения есть  $w_N(x, y)$ , то (19.11) означает, что

$$\lim_{N \rightarrow \infty} \iint (x - y)^2 w_N(x, y) dx dy = 0,$$

а (19.12) означает, что

$$\lim_{N \rightarrow \infty} \int dy \int_{y-\varepsilon}^{y+\varepsilon} w_N(x, y) dx = 1.$$

Так как  $\varepsilon$  сколь угодно мало, из последнего равенства вытекает, что при  $N \rightarrow \infty$   $w_N(x, y)$  переходит в  $w_{\xi}(y) \delta(x - y)$ .

Обратимся теперь к случайным *процессам*, т. е. к случайным функциям  $\xi(t)$  с непрерывным аргументом  $t$ .

В соответствии с тремя приведенными определениями вероятностной сходимости можно определить и непрерывность слу-

чайной функции  $\xi(t)$  в точке  $t$ . Речь идет о сходимости  $\xi(t + \Delta t)$  к  $\xi(t)$  при  $\Delta t \rightarrow 0$ . Следовательно, непрерывность в среднем квадратичном означает, что

$$\lim_{\Delta t \rightarrow 0} \overline{\{\xi(t + \Delta t) - \xi(t)\}^2} = 0,$$

или

$$\text{l.i.m. } \xi(t + \Delta t) = \xi(t).$$

Подчеркнем, что непрерывность случайной функции по вероятности отнюдь не предполагает непрерывности ее возможных значений. Последние могут быть и дискретными, как, например, заряд, поступивший на анод в задаче о дробовом шуме. Возможные значения заряда  $q(t)$  кратны заряду электрона  $e$ , т. е.  $q(t)$  меняется разрывно. Тем не менее функция  $q(t)$  предполагается непрерывной по вероятности, т. е. при  $\Delta t \rightarrow 0$

$$P\{|q(t + \Delta t) - q(t)| > e\} < \eta.$$

Другими словами, предполагается, что вероятность прихода на анод за время  $\Delta t$  заряда  $\Delta q = q(t + \Delta t) - q(t)$ , равного одному или нескольким  $e$ , стремится к нулю вместе с  $\Delta t$ .

Аналогично вероятностной непрерывности определяется и дифференцируемость случайной функции  $\xi(t)$ , а именно — как существование предела  $\dot{\xi}(t)$  (в каком-либо из указанных смыслов) выражения

$$\frac{\xi(t + \Delta t) - \xi(t)}{\Delta t}$$

при произвольном переходе  $\Delta t \rightarrow 0$ . Например, дифференцируемость в среднем квадратичном означает, что

$$\lim_{\Delta t \rightarrow 0} \overline{\left\{ \frac{\xi(t + \Delta t) - \xi(t)}{\Delta t} - \dot{\xi}(t) \right\}^2} = 0,$$

или

$$\text{l.i.m.}_{\Delta t \rightarrow 0} \frac{\xi(t + \Delta t) - \xi(t)}{\Delta t} = \dot{\xi}(t),$$

а дифференцируемость почти наверное означает, что

$$P\left\{ \lim_{\Delta t \rightarrow 0} \frac{\xi(t + \Delta t) - \xi(t)}{\Delta t} = \dot{\xi}(t) \right\} = 1.$$

Следует, однако, предостеречь от поспешного вывода, что переход от случайных последовательностей к случайным процессам представляет собой некий тривиальный шаг. В отношении процессов возможны такие, вполне оправданные, вопросы, ответы на которые связаны с определенными трудностями. Например, непрерывность (или дифференцируемость) процесса  $\xi(t)$  в некотором интервале значений  $t$  (т. е. в *несчетном* множестве

точек  $t$ ), вообще говоря, не принадлежит к случайным событиям, вероятности которых однозначно определяются конечными распределениями  $\xi(t)$ . Однако в соответствии с общими установками, принятыми в этой книге, мы не можем углубляться в такие вопросы, вновь отсылая читателя к математической литературе [2, 3].

В тех же трех смыслах сходимости можно понимать существование предела суммы

$$R_N = \sum_0^{N-1} \xi(t_i) (t_{i+1} - t_i),$$

где  $t_0 = a$ ,  $t_n = b$ , т. е. существование риманова интеграла от случайной функции  $\xi(t)$

$$R = \int_a^b \xi(t) dt.$$

Функция  $\xi(t)$  называется тогда  $R$ -интегрируемой (по вер., или п. н., или в ср. кв., смотря по тому, какая сходимость  $R_N$  к  $R$  имеется в виду). Аналогично можно определить интегралы Стильеса:

$$S = \int_a^b \xi(t) df(t), \quad S = \int_a^b f(t) d\xi(t),$$

где  $f(t)$  — детерминированная функция, — как пределы в каком-либо из вероятностных смыслов интегральных сумм

$$S_N = \sum_{i=0}^{N-1} \xi(t_i) [f(t_{i+1}) - f(t_i)],$$

$$S_N = \sum_{i=0}^{N-1} f(t_i) [\xi(t_{i+1}) - \xi(t_i)],$$

а также для двух независимых случайных функций  $\xi(t)$  и  $\eta(t)$  — интеграл Стильеса:

$$S = \int_a^b \xi(t) d\eta(t).$$

## § 20. Эргодичность случайного процесса

Свойство эргодичности, о котором далее пойдет речь, важно потому, что при его наличии имеют место чрезвычайно существенные соотношения *между функцией распределения и временем пребывания случайной функции  $\xi(t)$  в определенном интервале значений, между статистическими средними и средними по времени.*

Функции распределения, согласно «аксиоме измерения» вероятности, имеют *статистический* смысл. Это *относительные частоты в ансамбле одинаковых систем*, т. е. систем, в каждой из которых воспроизведены одни и те же условия протекания данного случайного процесса и одни и те же способы его регистрации или наблюдения. Если, например, речь идет о флуктуациях, то одинаковыми должны быть *макроскопические* характеристики всех систем, составляющих ансамбль. Имея ансамбль систем, мы располагаем обширным набором *реализаций* рассматриваемой случайной функции  $\xi(t)$  и с помощью соответствующих вероятностей, т. е. распределений систем ансамбля по возможным значениям  $\xi(t)$ , можем находить  $\overline{\xi(t)}$ ,  $\psi(t_1, t_2)$  и т. д.

Теоретики любят оперировать с ансамблями, но у экспериментаторов обычно одна лаборатория и одна установка, а не  $10^6$  или  $10^9$ <sup>1)</sup>. За данный промежуток времени  $(0, T)$  экспериментатор может получить лишь *одну* реализацию интересующего его случайного процесса и предпочитает поэтому усреднять *по времени*, пользуясь *одной* реализацией  $\xi(t)$ , одной осциллограммой или аналогичным образцом. Спрашивается, в каком соотношении находятся эти способы усреднения — по времени и по ансамблю?

Забегая вперед, укажем уже теперь, что для *стационарных* процессов, обладающих *свойством эргодичности*, оба способа усреднения при достаточно больших  $T$  практически совпадают, так как в этом случае стационарная вероятность состояния равна относительному времени пребывания системы в данном состоянии. Соответственно среднее статистическое равно среднему по достаточно большому промежутку времени<sup>2)</sup>. Говоря «равно» или «совпадает», мы, конечно, допускаем неточность, так как речь идет лишь о вероятностной сходимости — по вероятности, в среднем квадратичном или почти наверное.

Формулируем эти утверждения более точно и докажем их [4].

Пусть  $f(\xi)$  — некая детерминированная функция случайной функции  $\xi(t)$ . Обозначим ее среднее по времени за промежуток  $(0, T)$  индексом  $T$  и волнистой чертой сверху:

$$\tilde{f}_T(\xi(t)) = \frac{1}{T} \int_0^T f(\xi(t)) dt. \quad (20.1)$$

1) Конечно, случаи, когда ансамбль реализуется в одной установке, не исключены. Напомним, например, диффузию капли краски в воде. Правда, членами ансамбля здесь являются не сложные макроскопические системы, а малые частицы.

2) Именно поэтому стационарная вероятность имеет смысл вероятности состояния при термодинамическом равновесии, а средние статистические — смысл термодинамически равновесных значений.

Как сказано, это эмпирическая величина, получаемая в результате определенной обработки осциллограммы  $f(\xi)$  в интервале  $(0, T)$ . Вместе с тем это случайная величина, различная для разных реализаций  $\xi(t)$  в интервале  $(0, T)$ . Пусть  $w_1(x, t)$  и  $w_2(t, x; t', x')$  — одномерная и двумерная плотности вероятностей  $\xi(t)$ . Вычислим среднее значение и дисперсию случайной величины  $\overline{f_T(\xi)}$ . Имеем

$$\begin{aligned}\overline{f_T(\xi)} &= \frac{1}{T} \int_0^T \overline{f(\xi)} dt = \frac{1}{T} \int_0^T dt \int_{-\infty}^{+\infty} f(x) w_1(t, x) dx = \\ &= \int_{-\infty}^{+\infty} f(x) dx \frac{1}{T} \int_0^T w_1(t, x) dt = \int_{-\infty}^{+\infty} f(x) p_T(x) dx,\end{aligned}\quad (20.2)$$

где мы ввели «эффективную» плотность вероятности

$$p_T(x) = \frac{1}{T} \int_0^T w_1(t, x) dt. \quad (20.3)$$

Имеем, далее,

$$\begin{aligned}D[\overline{f_T(\xi)}] &= \langle \{\overline{f_T(\xi)} - \overline{f_T(\xi)}\}^2 \rangle = \\ &= \frac{1}{T^2} \int_0^T \int_0^T \{ \overline{f(\xi) f(\xi')} - \overline{f(\xi)} \overline{f(\xi')} \} dt dt' = \\ &= \frac{1}{T^2} \int_0^T \int_0^T \psi_f(t, t') dt dt',\end{aligned}\quad (20.4)$$

где  $\psi_f(t, t')$  — функция корреляции  $f(\xi)$ :

$$\begin{aligned}\psi_f(t, t') &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x) f(x') \{ w_2(t, x; t', x') - \\ &\quad - w_1(t, x) w_1(t', x') \} dx dx'.\end{aligned}\quad (20.5)$$

Таким образом, для всех  $f(\xi)$  с конечным  $\langle f^2(\xi) \rangle$  требование

$$\lim_{T \rightarrow \infty} \frac{1}{T^2} \int_0^T \int_0^T \psi_f(t, t') dt dt' = 0 \quad (20.6)$$

есть *необходимое и достаточное* условие того, чтобы имела место сходимость в среднем квадратичном:

$$\text{l.i.m.}_{T \rightarrow \infty} \{ \overline{f_T(\xi)} - \overline{f_T(\xi)} \} = 0, \quad (20.7)$$

причем необходимость вытекает просто из определения этого вида сходимости. В силу неравенства Чебышева для случайной величины  $\tilde{f}_T(\xi)$ :

$$P \{ |\tilde{f}_T(\xi) - \overline{\tilde{f}_T(\xi)}| \geq \varepsilon \} \leq \frac{D[\tilde{f}_T(\xi)]}{\varepsilon^2},$$

условие (20.6) достаточно и для сходимости по вероятности:

$$\lim_{T \rightarrow \infty} |\tilde{f}_T(\xi) - \overline{\tilde{f}_T(\xi)}| = 0 \text{ (по вер.)},$$

или в развернутой форме:

$$\lim_{T \rightarrow \infty} \left| \frac{1}{T} \int_0^T f[\xi(t)] dt - \int_{-\infty}^{+\infty} f(x) p_T(x) dx \right| = 0 \text{ (по вер.)},$$

где  $p_T(x)$  дается формулой (20.3). Можно сказать, что это закон больших чисел в применении к непрерывному наблюдению.

Докажем теперь, что при условии (20.6) *относительное время пребывания  $\xi(t)$  в промежутке  $(x, x+dx)$  сходится по вероятности при  $T \rightarrow \infty$  к  $p_\infty(x)dx$* , где

$$p_\infty(x) = \lim_{T \rightarrow \infty} p_T(x),$$

если, конечно, этот предел существует. Относительным временем пребывания называется отношение суммарного времени  $T_{(x, x+dx)}$ , проведенного  $\xi(t)$  в промежутке  $(x, x+dx)$ , т. е. суммы всех отмеченных на рис. 9 жирной линией отрезков оси абсцисс от момента 0 до момента  $T$ , к полной продолжительности интервала  $T$ .

Для доказательства достаточно взять в качестве  $f(\xi)$  функцию-фиксатор:

$$f(\xi) = \begin{cases} 1, & \text{если } x < \xi(t) \leq x + dx, \\ 0, & \text{если } \xi(t) \text{ вне } (x, x + dx). \end{cases}$$

Тогда по (20.1)

$$\tilde{f}_T(\xi) = \frac{T_{(x, x+dx)}}{T},$$

а по (20.2)

$$\overline{\tilde{f}_T(\xi)} = p_T(x) dx,$$

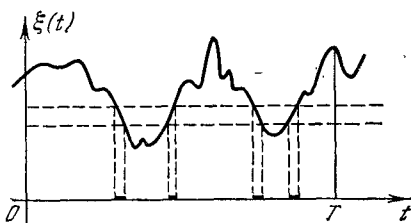


Рис. 9.

т. е. распределение  $p_T(x)dx$ , вычисляемое из одномерной плотности вероятности  $w_1(t, x)$  путем усреднения ее по параметру  $t$  [см. (20.3)], имеет смысл среднего (по ансамблю) относительного времени пребывания:

$$\frac{\overline{T(x, x+dx)}}{T} = p_T(x) dx = \frac{dx}{T} \int_0^T w_1(t, x) dt. \quad (20.8)$$

Теорема же (20.7) принимает теперь вид

$$\text{l.i.m.}_{T \rightarrow \infty} \left\{ \frac{T(x, x+dx)}{T} - p_T(x) dx \right\} = 0. \quad (20.9)$$

Все сказанное справедливо для случайного процесса  $f[\xi(t)]$ , удовлетворяющего условию (20.6) — так называемому *условию эргодичности* этого процесса. Необходимое и достаточное условие (20.6) можно заменить более сильными требованиями — *достаточными* условиями эргодичности, которые в приложениях теории большей частью оказываются выполненными. Например, условие (20.6) будет удовлетворено, если функция корреляции  $\psi_f(t, t')$  всюду ограничена и при всяком  $t$  убывает с увеличением  $|t' - t|$ , как  $|t' - t|^{-\alpha}$ , где  $\alpha > 0$ . В свою очередь, согласно (20.5), это будет выполнено, если при всяком  $t$  увеличение  $|t' - t|$  достаточно быстро приводит к статистической независимости между  $\xi(t)$  и  $\xi(t')$ :

$$w_2(t, x; t', x') \rightarrow w_1(t, x) w_1(t', x') \quad \text{при} \quad |t' - t| \rightarrow \infty. \quad (20.10)$$

Это последнее требование, означающее, в частности, достаточно быстрое «забывание» рассматриваемой системой ее предшествующих состояний, можно, согласно (15.2), формулировать и в терминах условной вероятности:

$$v(t', x' | t, x) \rightarrow w_1(t', x') \quad \text{при} \quad t' - t \rightarrow \infty. \quad (20.11)$$

Для марковского процесса это означает, что на вероятность перехода в состояние  $x'$  в момент  $t'$  достаточно давние прошедшие состояния уже не влияют и она превращается просто в одномерную вероятность состояния. Довольно часто под условиями эргодичности подразумевают именно требования (20.10) или (20.11), хотя при их невыполнении необходимое условие (20.6) может остаться в силе.

Приведем некоторые примеры распределения  $p_T(x)dx$ .

Пусть  $\xi(t) = F(t)\eta(t)$ , где  $F(t)$  — детерминированная функция времени, а  $\eta(t)$  — *стационарный* случайный процесс с одномерной функцией распределения  $w(y)dy$ . Одномерная плотность вероятности для  $\xi(t)$  может быть вычислена по известной



формуле, дающей распределение произведения двух случайных величин, если для  $F(t)$  тоже ввести «функцию распределения», а именно  $w_F(z) dz = \delta[z - F(t)] dz$ . Тогда

$$\begin{aligned} w_1(t, x) &= \int_{-\infty}^{+\infty} w\left(\frac{x}{y}\right) w_F(y) \frac{dy}{|y|} = \\ &= \int_{-\infty}^{+\infty} w\left(\frac{x}{y}\right) \delta[y - F(t)] \frac{dy}{|y|} = w\left[\frac{x}{F(t)}\right] \frac{1}{|F(t)|}. \end{aligned} \quad (20.12)$$

Подставляя это в (20.3), получаем для среднего относительного времени пребывания выражение

$$p_T(x) dx = \frac{dx}{T} \int_0^T w\left[\frac{x}{F(t)}\right] \frac{dt}{|F(t)|}. \quad (20.13)$$

Большую роль в радиотехнике играют так называемые *периодически-нестационарные процессы*, у которых одномерная плотность вероятности, а значит, и моменты  $\xi^n$  периодически зависят от  $t$ . Очевидно, в этом случае при вычислении  $p_T(x)$  целесообразно понимать под  $T$  период  $w_1(x, t)$ . Тем самым,  $p_T(x)$  не зависит от  $T$  и совпадает с  $p_\infty(x)$ .

Пусть  $w_1(t, x)$  — нормальный закон с гармонически меняющимся стандартом  $\sigma = \sigma_0(1 + k \cos \omega t)$ ,  $k \leq 1$ :

$$w_1(t, x) = \frac{1}{\sqrt{2\pi} \sigma_0 (1 + k \cos \omega t)} \exp \left\{ -\frac{x^2}{2\sigma_0^2 (1 + k \cos \omega t)^2} \right\}.$$

Вводя  $T = 2\pi/\omega$  и переменную интегрирования  $\varphi = \omega t$ , имеем

$$p_T(x) = \frac{1}{(2\pi)^{3/2} \sigma_0} \int_0^{2\pi} \exp \left\{ -\frac{x^2}{2\sigma_0^2 (1 + k \cos \varphi)^2} \right\} \frac{d\varphi}{1 + k \cos \varphi}.$$

Подстановка

$$\frac{1 - k^2}{1 + k \cos \varphi} = 1 - k \cos \theta$$

приводит этот интеграл к виду

$$p_T(x) = \frac{2}{(2\pi)^{3/2} \sigma_0 \sqrt{1 - k^2}} \int_0^\pi e^{-\beta (1 - k \cos \theta)^2} d\theta, \quad \beta = \frac{x^2}{2\sigma_0^2 (1 - k^2)^2}.$$

Если  $k \rightarrow 0$ , то  $w_1(t, x)$  переходит в стационарное гауссово распределение  $w_1(x)$  с дисперсией  $\sigma_0^2$ , и тогда  $p_T(x)$  совпадает с  $w_1(x)$ . Если же  $k \rightarrow 1$ , т. е.  $k = 1 - \varepsilon$ , где  $\varepsilon \rightarrow 0$ , то в интеграле существенны только малые значения  $\theta$ . Можно

приближенно положить

$$1 - k \cos \theta = \varepsilon + \theta^2/2 = \varepsilon (1 + s^2), \quad s = \theta/\sqrt{2\varepsilon}$$

и раздвинуть пределы интегрирования по  $s$  от 0 до  $\infty$ . Выражение для  $p_T(x)$  принимает вид

$$p_T(x) \approx \frac{2}{(2\pi)^{3/2} \sigma_0} \int_0^\infty \exp \left\{ -\frac{x^2}{8\sigma_0^2} (1 + s^2)^2 \right\} ds.$$

Это выражение, справедливое для достаточно малых  $\varepsilon$  (т.е.  $k \approx 1$ ), при  $x \gg \sigma_0$ , когда в показателе можно отбросить  $s^4$ , дает

$$p_T(x) \approx \frac{1}{\sqrt{2} \pi |x|} \exp \left\{ -\frac{x^2}{8\sigma_0^2} \right\} \quad (x \gg \sigma_0),$$

а в случае  $x \ll \sigma_0$ , когда в показателе существенно только  $s^4$ , дает

$$p_T(x) \approx \frac{2}{(2\pi)^{3/2} \sigma_0} \int_0^\infty \exp \left\{ -\frac{x^2 s^4}{8\sigma_0^2} \right\} ds = \frac{\sqrt[4]{2} \Gamma(1/4)}{4 \sqrt{\pi^3 \sigma_0} |x|} \quad (x \ll \sigma_0).$$

Мы видим отсюда, что даже кратковременное приближение стандарта  $\sigma$  к нулю влечет за собой значительное возрастание относительного времени пребывания  $\xi(t)$  в области малых значений  $x$ .

Как изменятся предыдущие результаты, если свойством эргодичности обладает *стационарный* процесс  $\xi(t)$ ? В этом случае  $\omega_1$  не зависит от  $t$ , а  $\omega_2$  зависит лишь от разности  $t' - t$ , в силу чего и функция корреляции для  $f(\xi)$  тоже зависит лишь от  $t' - t$ . Необходимое и достаточное условие эргодичности (20.6) принимает теперь следующий вид:

$$\lim_{T \rightarrow \infty} \frac{1}{T^2} \int_0^T \int_0^T \psi_f(t' - t) dt dt' = 0. \quad (20.14)$$

Отсюда нетрудно получить другую форму условия, содержащую только однократный интеграл от  $\psi_f(\tau)$ . С учетом четности  $\psi_f(t' - t)$  по аргументу  $\tau = t' - t$  имеем

$$\begin{aligned} \frac{1}{T^2} \iint_0^T \psi_f(t' - t) dt dt' &= \frac{2}{T^2} \int_0^T dt \int_0^t \psi_f(t' - t) dt' = \\ &= \frac{2}{T^2} \int_0^T dt \int_0^t \psi_f(\tau) d\tau. \end{aligned} \quad (20.15)$$

Но для функции корреляции стационарного в широком смысле процесса справедливо неравенство  $[\sigma^2 = \psi_f(0)]$

$$\frac{2\sigma^2}{T^2} \int_0^{2T} dt \int_0^t \psi_f(\tau) d\tau \geq \left[ \frac{2}{T} \int_0^T \psi_f(\tau) d\tau \right]^2. \quad (20.16)$$

(см. [5], а также задачу 3 гл. VI). В силу (20.15) и (20.16) выполнение условия (20.14) влечет за собой выполнение условия

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \psi_f(\tau) d\tau = 0, \quad (20.17)$$

которое является, таким образом, необходимым условием эргодичности стационарного случайного процесса. Но оно также и достаточно, поскольку (20.15) можно записать в виде

$$\frac{1}{T^2} \iint_0^T \psi_f(t' - t) dt dt' = \frac{2}{T^2} \int_0^T \frac{t}{T} dt \left( \frac{1}{t} \int_0^t \psi_f(\tau) d\tau \right),$$

откуда ясно, что при выполнении (20.17) имеет место и (20.14).

Условие (20.17), необходимое и достаточное для эргодичности стационарного в широком смысле случайного процесса, носит название условия Слущкого (см. [6], стр. 18). Оно допускает, что  $\psi_f(\tau)$  не стремится к нулю при  $\tau \rightarrow \infty$ , а, например, содержит члены вида  $a \cos \omega \tau$ <sup>1)</sup>. В этом случае

$$\frac{1}{T} \int_0^T \psi_f(\tau) d\tau = \frac{a \sin \omega T}{\omega T},$$

т. е. (20.17) выполнено.

Если интеграл от  $\psi_f(\tau)$  в пределах от 0 до  $\infty$  существует, то можно ввести эффективное время корреляции  $\tau_0$ :

$$\tau_0 = \frac{1}{\sigma_f^2} \int_0^\infty \psi_f(\tau) d\tau, \quad \sigma_f^2 = \psi_f(0)$$

(в действительности  $\psi_f(\tau)$  может обладать не одним, а несколькими характерными временными масштабами). В этом случае левая часть условия (20.17) при достаточно больших  $T$  (а именно  $T \gg \tau_0$ ) приближенно равна  $\sigma_f^2 \tau_0 / T$ .

<sup>1)</sup> Как мы видели (§ 18), функцией корреляции такого вида обладает гармоническое колебание со случайными и независимыми амплитудой и фазой, если, в частности, фаза распределена равномерно в интервале  $(0, 2\pi)$ .

*Достаточные* условия эргодичности стационарного процесса можно формулировать по-прежнему либо как требование

$$\lim_{\tau \rightarrow \infty} \psi_f(\tau) = 0, \quad (20.18)$$

либо аналогично (20.10) или (20.11):

$$\begin{aligned} \lim_{\tau \rightarrow \infty} w_2(x, \tau, x') &= w_1(x) w_1(x'), \\ \lim_{\tau \rightarrow \infty} v(x') | \tau, x | &= w_1(x'). \end{aligned} \quad (20.19)$$

Но, быть может, наиболее существенным обстоятельством является то, что для стационарного процесса распределение  $p_T(x)$  совпадает с одномерной функцией распределения  $w_1(x)$ , поскольку последняя не зависит от  $t$  [см. (20.3)]. Тем самым, теоремы (20.7)—(20.9) утверждают теперь, что при условии эргодичности (20.17) мы имеем

$$\text{l.i.m.}_{T \rightarrow \infty} \overline{f_T}(\xi) = \overline{f}(\xi), \quad (20.20)$$

т. е. среднее временное от  $f(\xi)$  сходится в среднем квадратичном (и по вероятности) к среднему статистическому и, в частности,

$$\text{l.i.m.}_{T \rightarrow \infty} \overline{\xi_T} = \text{l.i.m.}_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi(t) dt = \bar{\xi}. \quad (20.21)$$

Далее,

$$\text{l.i.m.}_{T \rightarrow \infty} \frac{T(x, x+dx)}{T} = w_1(x) dx, \quad (20.22)$$

т. е. *относительное время пребывания сходится в среднем квадратичном (и по вероятности) к одномерной функции распределения*. Последняя есть просто среднее статистическое от относительного времени пребывания:

$$\overline{\frac{T(x, x+dx)}{T}} = w_1(x) dx. \quad (20.23)$$

Это дает в руки экспериментатору, если он имеет дело со стационарным эргодическим случайным процессом, непосредственный метод измерения одномерной функции распределения: вероятность попадания  $\xi(t)$  в какой-либо интервал  $(a, b)$  измеряется относительным временем пребывания  $\xi(t)$  в этом интервале за достаточно длинный промежуток времени  $T$ .

Мы рассматривали до сих пор условия эргодичности применительно к временному среднему случайной функции  $f[\xi(t)]$ , т. е. функции, зависящей от значения  $\xi(t)$  в какой-то один мо-

мент времени (иногда это называют эргодичностью первого порядка); что можно сказать об эргодичности второго порядка, т. е. временных средних для функций  $f[\xi(t), \xi(t')]$ ? Ограничимся в этом вопросе частным случаем стационарного процесса  $\xi(t)$  и функцией  $f$ , равной произведению  $\xi(t)\xi(t')$ . Речь идет, следовательно, о том, при каком условии среднее временное

$$\overline{[\xi(t)\xi(t+\tau)]_T} = \frac{1}{T} \int_0^T \xi(t)\xi(t+\tau) dt$$

сходится при  $T \rightarrow \infty$  к своему среднему статистическому, равному, как легко видеть, функции корреляции процесса  $\xi(t)$ <sup>1)</sup>:

$$\overline{[\xi(t)\xi(t+\tau)]_T} = \frac{1}{T} \int_0^T \overline{\xi(t)\xi(t+\tau)} dt = \frac{1}{T} \int_0^T \psi(\tau) dt = \psi(\tau).$$

Очевидно, если такая сходимость имеет место, т. е.

$$\text{l.i.m.}_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi(t)\xi(t+\tau) dt = \psi(\tau), \quad (20.24)$$

то практически это означает возможность измерять функцию корреляции путем временного усреднения произведения  $\xi(t)\xi(t+\tau)$  по достаточно длинному промежутку времени  $T$ .

Равенство (20.24) означает, что дисперсия случайной величины  $[\xi(t)\xi(t+\tau)]_T$  стремится при  $T \rightarrow \infty$  к нулю. Если обозначить функцию корреляции случайной величины  $\eta_\tau(t) = \xi(t)\xi(t+\tau)$  через  $\psi_\eta(\sigma)$ :

$$\begin{aligned} \psi_\eta(\sigma) &= \langle \eta_\tau(t) \eta_\tau(t+\sigma) \rangle - \langle \eta_\tau(t) \rangle \langle \eta_\tau(t+\sigma) \rangle = \\ &= \langle \xi(t)\xi(t+\tau)\xi(t+\sigma)\xi(t+\sigma+\tau) \rangle - \psi^2(\tau), \end{aligned}$$

то, как показал Е. Е. Слуцкий, указанное требование сводится к тому, что

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \psi_\eta(\sigma) d\sigma = 0. \quad (20.25)$$

Если стационарность  $\xi(t)$  понимается в широком смысле (т. е. не утверждается, что функции распределения  $\omega_n$  для  $\xi(t)$  инвариантны по отношению к сдвигу начала отсчета времени при  $n \geq 3$ ), то независимость  $\psi_\eta$  от  $t$  представляет собой самостоятельное предположение. Если же стационарность  $\xi(t)$

<sup>1)</sup> Для простоты мы считаем, что  $\bar{\xi} = 0$ .

понимается в узком смысле, то момент четвертого порядка будет зависеть только от разностей моментов времени, т. е. от  $\tau$  и  $\sigma$ . Обычно следует из физических соображений, что зависимость между  $\eta(t)$  и  $\eta(t + \sigma)$  при  $\sigma \rightarrow \infty$  исчезает, так что  $\psi_\eta(\sigma) \rightarrow 0$ , и тогда условие (20.25) выполнено. Тем самым верна и теорема (20.24).

Наличие в  $\psi_\eta$  момента четвертого порядка не позволяет в общем случае выразить (20.25) — условие справедливости эргодической теоремы (20.24) для момента второго порядка — через функцию корреляции самого процесса  $\xi(t)$ , т. е. через  $\psi(\tau)$ . Однако в том частном и важном случае, когда  $\xi(t)$  — *нормальная* функция, а значит, *все* ее моменты, в том числе и четвертые, выражаются через  $\xi$  и  $\psi(\tau)$ , это возможно. А именно, если  $\xi(t)$  — нормальная функция (и для простоты  $\xi = 0$ ), то необходимым и достаточным условием для выполнения (20.24) будет

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T |\psi(\tau)|^2 d\tau = 0. \quad (20.26)$$

Если в функции корреляции  $\psi(\tau)$  содержатся периодические члены, то условие (20.25) для нее выполнено, но (20.26) уже не имеет места. Заметим, что (20.26) обеспечивает выполнение эргодической теоремы для моментов не только второго, но и любого порядка, т. е. все временные средние по промежутку  $(0, T)$  от произведений  $\xi(t + \tau_1)\xi(t + \tau_2) \dots \xi(t + \tau_n)$  сходятся по вероятности при  $T \rightarrow \infty$  к соответствующим моментам. Более того, для *любой функции* от *нормальных* величин  $\xi(t + \tau)$

$$f[\xi(t + \tau_1), \xi(t + \tau_2), \dots, \xi(t + \tau_n)]$$

(если среднее значение ее модуля  $|\overline{f}|$  конечно) теорема тоже верна: при  $T \rightarrow \infty$  имеем

$$\begin{aligned} \text{l.i.m.}_{T \rightarrow \infty} f_T &= \text{l.i.m.}_{T \rightarrow \infty} \frac{1}{T} \int_0^T f[\xi(t + \tau_1), \xi(t + \tau_2), \dots, \xi(t + \tau_n)] dt = \\ &= \overline{f(\xi_1, \xi_2, \dots, \xi_n)}. \end{aligned}$$

### Задачи

1. Пользуясь ответом задачи 5 гл. II, написать характеристическую функцию для  $n$ -мерного распределения нормального случайного процесса  $\xi(t)$  со средним значением  $\overline{\xi(t)} = 0$ .

Решение. Элементы корреляционной матрицы  $n$  случайных величин  $\xi(t_1), \dots, \xi(t_n)$  равны

$$b_{ik} = \overline{\xi(t_i) \xi(t_k)} = \psi(t_i, t_k),$$

так что

$$\varphi_n(u_1, t_1; u_2, t_2; \dots; u_n, t_n) = \exp \left\{ -\frac{1}{2} \sum_{i, k=1}^n \psi(t_i, t_k) u_i u_k \right\}.$$

Если процесс  $\xi(t)$  в широком смысле стационарен, то

$$\varphi_n(u_1, t_1; u_2, t_2; \dots; u_n, t_n) = \exp \left\{ -\frac{1}{2} \sum_{i, k=1}^n \psi(t_i - t_k) u_i u_k \right\},$$

а это означает, что и  $n$ -мерная плотность вероятности зависит при любом  $n$  только от разностей  $n$  моментов времени. Таким образом, для нормального процесса стационарность в широком смысле влечет за собой стационарность и в узком смысле.

2. Располагая двумерным распределением  $w_2$  случайного процесса  $\xi(t)$ , получить совместную функцию распределения  $\tilde{w}$  для  $\xi(t)$  и  $\xi(t+\tau)$ . Найти  $\tilde{w}$  для стационарного нормального процесса.

Решение. Имеем  $w_2(x, t; x_\tau, t+\tau)$ , где  $x_\tau$  — возможные значения  $\xi(t+\tau)$ . Введем вместо  $x_\tau$  новую переменную  $y = (x_\tau - x)/\tau$ , так что при  $\tau \rightarrow 0$  имеем  $y \rightarrow \dot{x}$ , например, в среднем квадратичном. С точностью до первого порядка по  $\tau$  можно написать  $x_\tau = x + y\tau$ , и, следовательно,

$$w_2(x, t; x_\tau, t+\tau) dx dx_\tau \approx w_2(x, t; x + y\tau, t+\tau) \tau dx dy.$$

В пределе при  $\tau \rightarrow 0$  получаем

$$\tilde{w}(x, y, t) dx dy = dx dy \lim_{\tau \rightarrow 0} \tau w_2(x, t; x + y\tau, t). \quad (1)$$

Для стационарного нормального процесса (для простоты с  $\overline{x_1^2} = \overline{x_2^2} = \sigma^2$ ) имеем

$$\varphi_2(u_1, u_2, \tau) = \exp \left\{ -\frac{\sigma^2}{2} (u_1^2 + u_2^2 + 2K(\tau) u_1 u_2) \right\},$$

где  $K(\tau)$  — коэффициент корреляции. Так как  $K(\tau)$  имеет при  $\tau = 0$  максимум (равный единице), степенное разложение  $K(\tau)$  есть

$$K(\tau) = 1 - \frac{1}{2} |K''|_0 \cdot \tau^2 + \dots$$

Следовательно,

$$\tilde{w}(x, y) =$$

$$= \lim_{\tau \rightarrow 0} \frac{\tau}{4\pi^2} \iint_{-\infty}^{+\infty} \exp \left\{ -\frac{\sigma^2}{2} \left[ u_1^2 + u_2^2 + 2 \left( 1 - \frac{1}{2} |K''|_0 \tau^2 \right) u_1 u_2 \right] - \right. \\ \left. - i \{ u_1 x + u_2 (x + y\tau) \} \right\} du_1 du_2.$$

Если ввести переменные интегрирования  $\alpha = u_1 + u_2$  и  $\beta = \frac{\tau}{2}(u_2 - u_1)$ , то получим

$$\begin{aligned} \tilde{w}(x, y) &= \lim_{\tau \rightarrow 0} \left\{ \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \exp \left[ -\frac{\sigma^2 \alpha^2}{2} \left( 1 - |K''|_0 \frac{\tau^2}{4} \right) + i\alpha \left( x + \frac{y\tau}{2} \right) \right] d\alpha \right\} \times \\ &\quad \times \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp \left\{ -\frac{\sigma^2 \beta^2}{2} |K''|_0 + i\beta y \right\} d\beta = \\ &= \frac{1}{\sqrt{2\pi} \sigma} \exp \left\{ -\frac{x^2}{2\sigma^2} \right\} \frac{1}{\sqrt{2\pi} \sigma \sqrt{|K''|_0}} \exp \left\{ -\frac{y^2}{2\sigma^2 |K''|_0} \right\}. \quad (2) \end{aligned}$$

Таким образом,  $\xi$  и  $\dot{\xi}$  нормальны и независимы (в один и тот же момент времени), причем

$$D[\xi] = \sigma^2, \quad D[\dot{\xi}] = \sigma^2 |K''(0)|. \quad (3)$$

3. Показать, что у стационарного в узком смысле случайного процесса  $\xi(t)$  смежные по порядку производные  $\xi^{(n)}(t)$  и  $\xi^{(n+1)}(t)$  некоррелированы.

Решение. Для вычисления  $n$ -й производной надо располагать  $(n+1)$ -мерным распределением самого процесса  $\xi(t)$ . Следовательно, стационарность и стационарная связанность *всех* производных требуют стационарности  $\xi(t)$  в узком смысле.

Для смежных по порядку производных, взятых в один и тот же момент времени  $t$ , имеем

$$\langle \xi^{(n)}(t) \xi^{(n+1)}(t) \rangle = \left\langle \frac{d}{dt} \left( \frac{1}{2} [\xi^{(n)}(t)]^2 \right) \right\rangle = \frac{1}{2} \frac{d}{dt} \langle [\xi^{(n)}(t)]^2 \rangle = 0,$$

так как средний квадрат любой производной — постоянная величина. В частности,

$$\overline{\xi \dot{\xi}} = 0, \quad \overline{\dot{\xi} \xi} = 0,$$

но  $\overline{\xi \ddot{\xi}}$ , вообще говоря, отлично от нуля.



## Г л а в а IV

### МАРКОВСКИЕ ПРОЦЕССЫ

#### § 21. Предварительные замечания

Хотя марковские процессы (процессы без вероятностного последствия) представляют собой весьма специальный класс случайных процессов, значение их очень велико, поскольку выделяющие их условия оказываются выполненными в широкой области приложений теории. Это тем более справедливо, что случайные процессы общего вида во многих случаях могут быть приведены к схеме процесса без последствия, если воспользоваться более детальным описанием рассматриваемого процесса, т. е. должным образом увеличить количество переменных, описывающих состояние рассматриваемой системы. Для пояснения этого обратимся вновь к случаю детерминированного процесса, но допустим, что динамическая система описывается дифференциальным уравнением не первого порядка (как это предполагалось в § 15), а второго. Тогда решение при начальных условиях

$$x = x_0, \quad \dot{x} = u_0 \quad \text{при} \quad t = t_0$$

будет

$$x = f(t, x_0, u_0, t_0).$$

Если и теперь понимать под состоянием системы только координату  $x$ , то для плотности условной вероятности значения  $x$  в момент  $t$  надо было бы написать

$$v(t, x | t_0, x_0, u_0) = \delta[x - f(t, x_0, u_0, t_0)]. \quad (21.1)$$

Но задание  $u_0$  равносильно заданию *двух* значений  $x$  в весьма близкие моменты времени [скажем,  $x_0$  при  $t = t_0$  и  $x_0 - h$  при  $t = t_0 - \tau$ , так что  $u_0 = \lim_{\tau \rightarrow 0} (h/\tau)$ ]. Таким образом, условная

вероятность (21.1), по существу, зависит от *двух* предшествующих состояний:

$$v(t, x | t_0, x_0, u_0) \approx v(t, x | t_0, x_0; t_0 - \tau, x_0 - h),$$

и, следовательно, статистическое обобщение (21.1) не является процессом марковского типа.

Не только в статистической, но и в динамической теории обычно предпочитают избегать зависимости состояния системы от ее поведения до фиксированного начального момента. Это достигается *расширением самого понятия состояния* в момент  $t$  путем введения новых характеризующих состояние величин. В приведенном примере применение этого приема сводится к тому, что наряду с координатой  $x$  вводится еще и скорость  $\dot{x} = u$ . Понимая под состоянием совместное задание  $x$  и  $u$  в момент  $t$ , можно записать условную вероятность этого состояния для рассматриваемого динамического процесса в виде

$$\delta[x - f(t, x_0, u_0, t_0)] \delta[u - \dot{f}(t, x_0, u_0, t_0)],$$

что представляет собой частный случай вероятности перехода  $v(t, x, u | t_0, x_0, u_0)$ . Таким образом, в соответствующей статистической схеме мы приходим теперь к *марковскому* процессу, но для совокупности *двух* случайных функций  $x$  и  $u$ , т. е. для двумерной случайной функции.

Аналогичным образом  $k$ -мерный случайный процесс, не являющийся марковским, можно путем введения достаточно обширной совокупности «координат» сделать марковским, но для более высокого числа измерений  $k'$ . Грубо говоря, для этого достаточно понимать под «состоянием» системы совокупность значения рассматриваемого процесса в последний наблюдаемый момент времени  $t$  и некоторого количества значений из «предыстории» этого процесса при  $t' < t$ . Если, однако, требовать конечности  $k'$ , то этот прием будет осуществим не всегда. Именно так обстоит дело в приведенном ранее примере движения частицы в жидкости, оказывающей вязкое последствие (§ 25). Последнее описывается интегро-дифференциальным уравнением, т. е. не может быть исчерпано никакой конечной совокупностью производных от  $x$  в момент  $t$  или конечной совокупностью значений  $x$  в моменты  $t' \leq t$ .

В зависимости от дискретности или непрерывности возможных значений параметра  $t$  и возможных значений  $x$  самой случайной функции  $\xi(t)$  можно и для случайных функций марковского типа различать те же четыре разновидности, о которых говорилось в § 14.

Если время  $t$  дискретно, то вместо значений  $t_n$  можно рассматривать в качестве аргумента номера  $n$  этих значений и го-

ворить об «испытаниях», занумерованных целыми числами  $n$ . Случайная функция  $\xi(t)$  сводится тогда к последовательности случайных величин  $\xi(t_n) \equiv \xi_n$  ( $\xi$  при  $n$ -м испытании). Хотя схемы с дискретно меняющимся параметром  $t$  в некоторых отношениях более просты и наглядны, мы сосредоточим внимание на собственно процессах, т. е. на случае непрерывного изменения  $t$ , который впервые был рассмотрен в работах Эйнштейна и Смолуховского по теории брауновского движения. В отношении марковских последовательностей мы ограничимся отдельными замечаниями.

Наиболее близкими к классической теории вероятностей являются *дискретные последовательности*, т. е. процессы с *прерывным временем* и с дискретными возможными значениями случайных величин  $\xi_n = \xi(t_n)$ . Обозначим эти возможные значения через  $x_i$  ( $i = 1, 2, \dots$ ). Марковость процесса означает, что существует вероятность перехода от любого из значений  $x_i$  при  $l$ -м испытании к любому значению  $x_k$  при  $n$ -м испытании ( $n > l$ ):

$$p(n, x_k | l, x_i) = P\{\xi_n = x_k | \xi_l = x_i\}. \quad (21.2)$$

Если, в частности, число возможных значений (состояний)  $x_i$  конечно ( $i = 1, 2, \dots, N$ ), то процесс называется *простой цепью Маркова*.

Однородность рассматриваемого процесса марковского типа состоит в том, что вероятность перехода зависит от  $n$  и  $l$  только через их разность, т. е. зависит только от числа «шагов»  $s = n - l$ , пройденных от начального испытания к конечному.

Очевидно, последовательности марковского типа можно рассматривать как непосредственное обобщение последовательностей независимых испытаний, о которых говорилось в гл. I. В терминах вероятностей перехода можно сказать, что у последовательности независимых испытаний  $p(n, x_k | l, x_i) = P\{\xi_n = x_k\}$ , т. е. вероятность перехода в состояние  $x_k$  просто совпадает с вероятностью этого состояния при  $n$ -м испытании, независимо от результатов других испытаний.

Если случайные величины  $\xi_n$  имеют непрерывное множество возможных значений  $x$ , то марковость такого *непрерывного процесса с дискретным временем* означает существование вероятности перехода, которую можно записать как в интегральной форме:

$$V(n, x | l, y) = P\{\xi_n \leq x | \xi_l = y\},$$

так и в дифференциальной:

$$v(n, x | l, y) dx = d_x V = P\{x < \xi_n \leq x + dx | \xi_l = y\}.$$

## § 22. Уравнение Смолуховского

Фундаментальным для теории случайных функций марковского типа является уравнение Смолуховского [1], выражающее тот простой факт, что вероятности перехода для каких-либо *трех* последовательных моментов времени должны быть определенным образом согласованы между собой. Если взять для конкретности марковский процесс ( $t$  непрерывно) и непрерывные же возможные значения случайной функции  $\xi(t)$ , то речь идет о следующем.

Переход из состояния  $x_0$  в момент времени  $t_0$  в состояние  $x$  в момент  $t$  [вероятность этого перехода есть  $v(t, x | t_0, x_0) dx$ ] реализуется через какое-либо состояние  $y$  в промежуточный момент времени  $\theta$  ( $t_0 < \theta < t$ ), т. е. может быть представлен в виде двух последовательных переходов  $x_0 \rightarrow y$  и  $y \rightarrow x$ . Вероятность совместного осуществления этих двух переходов есть

$$v(\theta, y | t_0, x_0) dy v(t, x | \theta, y) dx,$$

причем с теми же самыми условными плотностями вероятности  $v$  (марковский процесс). Но переход  $x_0 \rightarrow y \rightarrow x$  есть один из взаимно исключающих частных случаев перехода  $x_0 \rightarrow x$

(рис. 10), так что суммирование по всем возможным состояниям  $y$  в момент  $\theta$  должно давать полную вероятность перехода  $x_0 \rightarrow x$ :

$$\begin{aligned} v(t, x | t_0, x_0) dx &= \\ &= dx \int v(t, x | \theta, y) v(\theta, y | t_0, x_0) dy. \end{aligned} \quad (22.1)$$

Это и есть уравнение Смолуховского.

Следует вновь обратить внимание на существенное ограничение, которое налагается уравнением (22.1) на допустимый вид функций  $v$ : интегрирование по  $y$  произведения двух  $v$  должно снова давать функцию  $v$  и должно автоматически исключать зависимость результата от промежуточного момента времени  $\theta$ , от которого зависит подынтегральная функция. По существу, это лишь иное выражение тех ограничений вида  $n$ -мерных функций распределения, которыми выделены марковские процессы. Действительно, нетрудно получить уравнение Смолуховского и другим путем, опираясь только на условие согласованности распределений различной мерности (§ 14) и на определение марковости (15.3) (см. задачу 1).

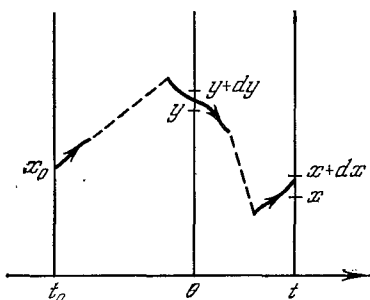


Рис. 10.

Проинтегрировав (22.1) по  $x$  от нижней границы интервала возможных значений до  $x$  и перейдя от риманова интеграла по  $y$  к интегралу Стильеса, мы получаем уравнение Смолуховского для *интегральных* вероятностей перехода:

$$V(t, x | t_0, x_0) = \int V(t, x | \theta, y) d_y V(\theta, y | t_0, x_0) \quad (t_0 < \theta < t). \quad (22.2)$$

В таком виде оно охватывает случаи как непрерывных, так и дискретных возможных значений  $\xi(t)$ . Это различие, в общем, является второстепенным, так как, даже не прибегая к интегралу Стильеса, а оперируя плотностями вероятности, мы всегда можем при дискретных возможных значениях записывать эти плотности через соответствующие дельта-функции. Приведем все же запись уравнения Смолуховского для процесса с дискретными возможными значениями случайной функции  $\xi(t)$ , т. е. для вероятностей перехода  $p(t, x_k | t_0, x_i)$ :

$$p(t, x_k | t_0, x_i) = \sum_j p(t, x_k | \theta, x_j) p(\theta, x_j | t_0, x_i) \quad (t_0 < \theta < t). \quad (22.3)$$

Если речь идет о случайных марковских *последовательностях*, т. е. значения параметра  $t$  дискретны, то в уравнениях (22.1) — (22.3) можно заменить аргументы  $t_n$  на «номера испытаний»  $n$ . Например, вместо (22.3) получим

$$p(n, x_k | l, x_i) = \sum_j p(n, x_k | m, x_j) p(m, x_j | l, x_i) \quad (l < m < n). \quad (22.4)$$

Для случая, когда число возможных дискретных значений *конечно* (цепи Маркова), т. е.  $i, j, k = 1, 2, \dots, N$ , уравнение (22.4) было установлено Марковым и носит название *уравнения Маркова*.

Если случайная функция марковского типа *однородна по  $t$* , то (22.1) примет вид

$$v(x | t - t_0, x_0) = \int_{-\infty}^{+\infty} v(x | t - \theta, y) v(y | \theta - t_0, x_0) dy \quad (t_0 < \theta < t), \quad (22.5)$$

а вместо (22.4) будет

$$p(x_k | n - l, x_i) = \sum_j p(x_k | n - m, x_j) p(x_j | m - l, x_i) \quad (l < m < n). \quad (22.6)$$

Как сказано, мы не будем углубляться в теорию марковских последовательностей и цепей (см. [2]), но все же рассмотрим в качестве примера применение (22.6) к однородной цепи Маркова. Введем следующее более лаконичное обозначение для вероятности перехода за  $s$  шагов:

$$p(x_k | s, x_i) = p_{ik}(s).$$

Уравнение Маркова (22.6) запишется тогда в виде

$$p_{ik}(s) = \sum_{j=1}^N p_{ij}(m) p_{jk}(s-m) \quad (0 < m < s). \quad (22.7)$$

Величины  $p_{ik}(s)$  образуют матрицу  $\pi_s$  вероятностей перехода (коротко — матрицу перехода) за  $s$  шагов. Уравнение (22.7) показывает, что матрица  $\pi_s$  равна произведению матриц  $\pi_m$  и  $\pi_{s-m}$ :

$$\pi_s = \pi_m \pi_{s-m} \quad (0 < m < s).$$

Но при  $s=2$  возможно только  $m=1$ , так что  $\pi_2 = \pi_1^2$ . При  $s=3$  имеем  $\pi_3 = \pi_1 \pi_2 = \pi_2 \pi_1 = \pi_1^3$ , и, вообще,

$$\pi_s = \pi_1^s. \quad (22.8)$$

По этой формуле вероятности перехода за  $s$  шагов выражаются через вероятности перехода  $p_{ik}(1) \equiv p_{ik}$  за один шаг. Заметим, что в матрице перехода

$$\pi_1 = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ p_{N1} & p_{N2} & \cdots & p_{NN} \end{pmatrix}$$

все элементы, разумеется, неотрицательны, причем ни один столбец не состоит сплошь из нулей (в числе  $x_k$  содержатся только *возможные* конечные состояния), равно как и ни одна строка не состоит из нулей, поскольку при всяком начальном состоянии  $x_i$



Рис. 11.

$$\sum_{k=1}^N p_{ik} = 1.$$

Пусть, например, рассматриваются случайные блуждания частицы между двумя отражающими стенками (рис. 11). Всюду, кроме положений  $i=1$  и  $i=N$ , шаг вправо имеет вероятность  $p$ , шаг влево — вероятность  $q=1-p$ . На стенках частица обязательно отступает на один шаг внутрь. Это типичная однородная цепь Маркова со следующими вероятностями

перехода за один шаг ( $1 < i < N$ ):

$$p_{ik} = \begin{cases} p & (k = i + 1), \\ q & (k = i - 1), \\ 0 & (\text{остальные } k), \end{cases}$$

$$p_{1k} = \begin{cases} 1 & (k = 2), \\ 0 & (\text{остальные } k), \end{cases} \quad p_{Nk} = \begin{cases} 1 & (k = N - 1), \\ 0 & (\text{остальные } k). \end{cases}$$

Таким образом, матрица  $\pi_1$  имеет вид

$$\pi_1 = (p_{ik}) = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ q & 0 & p & 0 & \dots & 0 & 0 & 0 \\ 0 & q & 0 & p & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & q & 0 & p \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 & 0 \end{pmatrix}.$$

Нетрудно убедиться в том, что и для *неоднородной* марковской последовательности вероятности перехода за  $s$  шагов тоже могут быть выражены через вероятности перехода за один шаг. Последние зависят теперь от номера испытания, так что мы обозначим их следующим образом:

$$p(n+1, y|n, x) = p_n(y|x).$$

Согласно (22.4)

$$p(n+2, y|n, x) = \sum_z p_{n+1}(y|z) p_n(z|x),$$

$$\dots$$

$$p(n+s, y|n, x) = \sum_{z, u, \dots, v} p_{n+s-1}(y|z) p_{n+s-2}(z|u) \dots p_n(v|x).$$

Для непрерывной функции  $\xi(t)$  тот же результат можно записать через плотности вероятности:

$$v(n+s, y|n, x) =$$

$$= \int \dots \int v_{n+s-1}(y|z) v_{n+s-2}(z|u) \dots v_n(w|x) dz du \dots dw.$$

Нетрудно сообразить, что обе формулы представляют собой просто повторные композиции соответствующих условных функций распределения, так что их можно записать в виде

$$p(n+s, y|n, x) = p_{n+s-1}(y|z) * p_{n+s-2}(z|u) * \dots * p_n(w|x).$$

В частном случае однородной цепи Маркова, когда вероятности при однократном переходе не зависят от его номера, это дает для матрицы перехода за  $s$  шагов формулу (22.8).

### § 23. Марковский процесс с дискретными состояниями

Перепишем уравнение Смолуховского (22.3) для трех последовательных моментов времени  $t_0 < t < t + \tau$ :

$$p(t + \tau, x_k | t_0, x_i) = \sum_j p(t + \tau, x_k | t, x_j) p(t, x_j | t_0, x_i). \quad (23.1)$$

Допустим, что для достаточно малых  $\tau$  вероятность перехода имеет вид

$$p(t + \tau, x_k | t, x_j) = \delta_{jk} + A_{jk}\tau + o(\tau), \quad (23.2)$$

где  $o(\tau)$  — члены выше первого порядка относительно  $\tau$ . Мы учитываем в (23.2), что при  $\tau \rightarrow 0$  конечное состояние с достоверностью совпадает с начальным, т. е.

$$p(t, x_k | t, x_j) = \delta_{jk}, \quad (23.3)$$

а кроме того, предполагаем существование производной

$$A_{jk}(t) = \left. \frac{\partial p(u, x_k | t, x_j)}{\partial u} \right|_{u=t}. \quad (23.4)$$

Очевидно, при  $j \neq k$  (т. е.  $\delta_{jk} = 0$ )  $A_{jk}(t)dt$  есть вероятность перехода за время  $dt$  из состояния  $x_j$  в другое состояние  $x_k$ . Тем самым, при  $j \neq k$  должно быть

$$A_{jk}(t) \geq 0 \quad (j \neq k). \quad (23.5)$$

Так как вероятность перехода в *какое-нибудь* из возможных состояний должна равняться единице:

$$\sum_k p(t + \tau, x_k | t, x_j) = 1, \quad (23.6)$$

имеем

$$\sum_k A_{jk}(t) = 0, \quad A_{jj}(t) = - \sum_{k \neq j} A_{jk}(t) \leq 0. \quad (23.7)$$

Подставим (23.2) в правую часть (23.1) и перейдем к пределу при  $\tau \rightarrow 0$ . Это приводит к следующему основному уравнению (точнее — системе уравнений) для марковского процесса с дискретными состояниями<sup>1)</sup>:

$$\frac{\partial p(t, x_k | t_0, x_i)}{\partial t} = \sum_j A_{jk}(t) p(t, x_j | t_0, x_i). \quad (23.8)$$

При начальных условиях (23.3) эти уравнения определяют зависимость вероятностей перехода от времени. Если число возможных состояний конечно, то нетрудно показать, что для

<sup>1)</sup> Строгий вывод см. у А. Н. Колмогорова [3]. Заметим, что в силу (23.3) и (23.4) первоначальное предположение  $t > t_0$  можно отбросить, т. е. в (23.8) допустимо и  $t = t_0$ .



любых непрерывных  $A_{ik}(t)$ , удовлетворяющих условиям (23.5) и (23.7), уравнение (23.8) с начальным условием (23.3) имеет единственное решение, неотрицательное и удовлетворяющее (23.1) и (23.6), т. е. определяет марковский процесс. При бесконечном множестве возможных состояний вопрос о единственности и допустимости решения требует специального исследования, которое и проводилось в ряде математических работ. С точки зрения практических приложений этот вопрос не является обычно особо острым.

Наряду с (23.8) справедливы также следующие дифференциальные уравнения по начальному моменту  $t_0$  [3]:

$$\frac{\partial p(t, x_k | t_0, x_i)}{\partial t_0} = - \sum_j A_{ij}(t_0) p(t, x_k | t_0, x_j), \quad (23.9)$$

которые легко получить, выбрав промежуточный момент времени близким не к конечному, а к начальному моменту  $t_0$ .

Уравнениям (23.8) удовлетворяют не только вероятности перехода, но и одномерная вероятность состояния  $p_1(t, x_k)$ . Действительно, если задана начальная функция распределения  $P(t_0, x_i)$ , то

$$p_1(t, x_k) = \sum_i P(t_0, x_i) p(t, x_k | t_0, x_i). \quad (23.10)$$

Умножив (23.8) на  $P(t_0, x_i)$  и взяв сумму по  $i$ , получим в силу (23.10)

$$\frac{\partial p_1(t, x_k)}{\partial t} = \sum_j A_{jk}(t) p_1(t, x_j). \quad (23.11)$$

Эти уравнения надо интегрировать при начальных условиях

$$p_1(t_0, x_k) = P(t_0, x_k). \quad (23.12)$$

Если рассматриваемый марковский процесс однороден, т. е. вероятность перехода имеет вид  $p(x_k | t, x_i)$  ( $t$  вместо  $t - t_0$ ), то, согласно (23.4),  $A_{jk}$  — постоянные величины и уравнения (23.8) принимают вид

$$\frac{\partial p(x_k | t, x_i)}{\partial t} = \sum_j A_{jk} p(x_j | t, x_i). \quad (23.13)$$

В предположении, что при  $t \rightarrow \infty$  вероятности перехода «забывают» об исходном состоянии и превращаются в стационарные вероятности состояния (наличие эргодичности)<sup>1)</sup>:

$$\lim_{t \rightarrow \infty} p(x_k | t, x_i) = P(x_k),$$

<sup>1)</sup> Это будет, в частности, в том случае, если все  $A_{jk} \neq 0$ , т. е. все переходы при малых  $t$  возможны.

мы получаем для нахождения  $P(x_k)$ , согласно (23.6) и (23.13), уравнения<sup>1)</sup>

$$\sum_k P(x_k) = 1, \quad \sum_j A_{jk} P(x_j) = 0. \quad (23.14)$$

Рассмотрим два примера, относящиеся к *однородным* процессам.

1. *Двусторонняя реакция*. Система может при этом находиться в одном из двух состояний 1 и 2. В частности, речь может идти о распаде частицы (ионизация, диссоциация, химическое разложение, радиоактивный распад), когда состояние 1 — это нераспавшаяся частица, а состояние 2 — распавшаяся.

В общем случае возможен как процесс распада (переход  $1 \rightarrow 2$  с вероятностью  $\alpha dt$  за время  $dt$ ), так и процесс восстановления (переход  $2 \rightarrow 1$  с вероятностью  $\beta dt$  за время  $dt$ ). Таким образом,  $A_{12} = \alpha$ ,  $A_{21} = \beta$  и, согласно (23.7),  $A_{11} = -\alpha$ ,  $A_{22} = -\beta$ , так что уравнения (23.11) для вероятностей состояния

$$p_1(t, 1) \equiv P_1(t) \quad \text{и} \quad p_1(t, 2) \equiv P_2(t)$$

принимают вид

$$dP_1/dt = -\alpha P_1 + \beta P_2, \quad dP_2/dt = \alpha P_1 - \beta P_2.$$

Второго уравнения можно было бы и не писать, а исключить  $P_2$  из первого, пользуясь тем, что  $P_1 + P_2 = 1$ . Тогда

$$dP_1/dt = -\lambda P_1 + \beta \quad (\lambda = \alpha + \beta).$$

Пусть при  $t = 0$  задано  $P_1 = 1$  (соответственно  $P_2 = 0$ ), т. е. система с достоверностью находится первоначально в состоянии 1. Решение будет

$$P_1(t) = e^{-\lambda t} + \frac{\beta}{\lambda} (1 - e^{-\lambda t}), \quad P_2(t) = \frac{\alpha}{\lambda} (1 - e^{-\lambda t}).$$

При  $t \rightarrow \infty$  достигаются стационарные значения обеих вероятностей, не зависящие от начальных условий:

$$P_1 = \beta/\lambda, \quad P_2 = \alpha/\lambda,$$

т. е. процесс эргодичен. Если восстановление невозможно ( $\beta = 0$ ), как это имеет место, например, в радиоактивном распаде, то

$$P_1(t) = e^{-\lambda t}, \quad P_2(t) = 1 - e^{-\lambda t}.$$

2. *Двухпозиционное реле*. Пусть оно находится под воздействием случайной последовательности управляющих импульсов,

<sup>1)</sup> Предполагается, что  $\partial p/\partial t \rightarrow 0$  при  $t \rightarrow \infty$ , что, вообще, не следует из существования постоянных пределов у самих  $p$ .

с одинаковой вероятностью имеющих знаки плюс или минус. Пусть положительный импульс создает или сохраняет состояние 1, отрицательный — создает или сохраняет состояние 2. Тогда  $A_{12} = A_{21} = \alpha$  (т. е.  $\alpha dt$  — вероятность *изменения* состояния за время  $dt$ ) и, согласно (23.7),  $A_{11} = A_{22} = -\alpha$ . Очевидно, мы имеем здесь частный случай предыдущей схемы, соответствующий  $\beta = \alpha$ . Но теперь мы напомним уравнения (23.13) для четырех *вероятностей перехода* ( $x = 1, 2$ ):

$$dp(1|t, x)/dt = -\alpha p(1|t, x) + \alpha p(2|t, x),$$

$$dp(2|t, x)/dt = \alpha p(1|t, x) - \alpha p(2|t, x),$$

которые надо решать при начальных условиях (23.3). Решения имеют вид

$$p(2|t, 1) = p(1|t, 2) = \frac{1}{2}(1 - e^{-2\alpha t}),$$

$$p(1|t, 1) = p(2|t, 2) = \frac{1}{2}(1 + e^{-2\alpha t}).$$

Мы снова видим, что процесс эргодичен: при  $t \rightarrow \infty$  вероятности перехода стремятся к стационарным значениям (как это и должно быть, поскольку все  $A_{jk} \neq 0$ ), а именно:

$$P(1) = P(2) = \frac{1}{2}.$$

В обоих примерах мы имеем монотонное приближение вероятностей к их предельным значениям. Возможен, однако, и осцилляторный ход, т. е. затухающие колебания. Если искать частное решение уравнений (23.13), пропорциональное  $e^{\lambda t}$ , то для определения  $\lambda$  в случае  $N$  возможных значений получится характеристическое уравнение

$$\begin{vmatrix} A_{11} - \lambda & A_{21} & \dots & A_{N1} \\ A_{12} & A_{22} - \lambda & \dots & A_{N2} \\ \dots & \dots & \dots & \dots \\ A_{1N} & A_{2N} & \dots & A_{NN} - \lambda \end{vmatrix} = 0,$$

т. е. уравнение  $N$ -й степени относительно  $\lambda$ . Эргодичность будет иметь место в том случае, если полином по степеням  $\lambda$ , остающийся в левой части после сокращения на некоторую степень  $\lambda$ , будет полиномом Гурвица, т. е. все отличные от нуля корни  $\lambda$  будут иметь отрицательные вещественные части. Очевидно, затухающие колебания будут происходить в том случае, когда среди корней имеются комплексно-сопряженные пары.

В примере двухпозиционного реле имелось два корня, а именно  $\lambda = 0$  и  $\lambda = -2\alpha$ . В цитированной работе А. Н. Колмогорова [3] приведен следующий пример для  $N = 3$ :

$$A_{12} = A_{23} = A_{31} = \alpha, \quad A_{21} = A_{32} = A_{13} = 0, \quad A_{11} = A_{22} = A_{33} = -\alpha.$$

Корни характеристического уравнения здесь равны

$$\lambda = 0, \quad \lambda = -\frac{3a}{2} + i \frac{\sqrt{3}a}{2}, \quad \lambda = -\frac{3a}{2} - i \frac{\sqrt{3}a}{2},$$

так что вероятности перехода стремятся к стационарным значениям  $[P(1) = P(2) = P(3) = 1/3]$ , совершая затухающие колебания<sup>1)</sup>.

#### § 24. Переход от дискретной последовательности к процессу с непрерывным множеством состояний. Распределение Релея

Прежде чем вывести основное дифференциальное уравнение для марковских процессов с непрерывным множеством возможных состояний, укажем на применяемый иногда упрощенный прием: время  $t$  и множество состояний  $x$  подразделяются на весьма малые промежутки  $\Delta t$  и  $\Delta x$ , составляется уравнение для марковской последовательности, а затем делается переход к пределу при  $\Delta t \rightarrow 0$  и  $\Delta x \rightarrow 0$ .

Рассмотрим конкретный пример блуждания частицы, которая в моменты времени  $N \Delta t$  ( $N = 1, 2, \dots$ ) скачет на шаг  $\Delta x = a$  вправо или влево с вероятностями  $p$  и  $q = 1 - p$ . Это уже знакомая задача об «абсолютно пьяном человеке» или о сложении колебаний со случайными амплитудами  $\pm a$  (§§ 4 и 6). Мы имеем здесь однородную марковскую последовательность, но не стационарную (распределение зависит от номера испытания).

Напомним, что вероятность того, что за  $N$  шагов будет сделано  $n$  шагов вправо [частица уйдет вправо на  $x = (2n - N)a$ , или же интенсивность суммы  $N$  колебаний будет  $J = (2n - N)^2 a^2$ ], дается биномиальным законом  $P_N(n) = C_N^n p^n q^{N-n}$ . Но сейчас мы подойдем к этой задаче с несколько иной стороны.

Обозначим абсциссу частицы в момент  $t = N \Delta t$  через  $x = ta$  и, считая, что в момент  $t = 0$  частица вышла из точки  $x = 0$ , введем *вероятность перехода* за  $N$  шагов:

$$p(N \Delta t, ta | 0, 0) \equiv p_{N,m}.$$

Нетрудно связать  $p_{N,m}$  с биномиальной вероятностью  $P_N(n)$ , но, как сказано, теперь мы будем искать  $p_{N,m}$ , не обращаясь к  $P_N(n)$  (см. задачу 6).

Частица может попасть при  $N$ -м шаге в точку  $x = ta$ , если при  $(N-1)$ -м шаге она оказалась в  $(m-1)a$  или в  $(m+1)a$ , причем из первого положения — с вероятностью  $p$ , а из вто-

<sup>1)</sup> Другие примеры и обобщение на многомерный случай см. в [4], § 3.2 и [5].

рого — с вероятностью  $q$ . Следовательно, по формуле полной вероятности имеем

$$p_{N,m} = p p_{N-1,m-1} + q p_{N-1,m+1}$$

— уравнение Маркова для данной задачи. Если изменить обозначение вероятности перехода, введя  $v(t, x)a = p_{N,m}$ , то уравнение переписывается в виде

$$v(t, x) = p v(t - \Delta t, x - a) + q v(t - \Delta t, x + a).$$

Разложим теперь  $v$  в правой части по степеням  $\Delta t$  и  $a = \Delta x$ . Учитывая, что  $p + q = 1$ , получаем

$$\begin{aligned} 0 = & -\frac{\partial v}{\partial t} \Delta t - (p - q) \frac{\partial v}{\partial x} a + \frac{\partial^2 v}{\partial x^2} \frac{a^2}{2} + \\ & + (p - q) \frac{\partial^2 v}{\partial x \partial t} a \Delta t + \frac{\partial^2 v}{\partial t^2} \frac{(\Delta t)^2}{2} + \dots \end{aligned} \quad (24.1)$$

Разделим уравнение на  $\Delta t$  и перейдем к пределу при  $\Delta t \rightarrow 0$ ,  $p - q \rightarrow 0$ ,  $a \rightarrow 0$ , причем предположим, что существуют следующие конечные пределы:

$$\lim_{\Delta t \rightarrow 0} \frac{(p - q)a}{\Delta t} = A, \quad \lim_{\Delta t \rightarrow 0} \frac{a^2}{\Delta t} = B. \quad (24.2)$$

Таким образом, допускается, что  $a \sim (\Delta t)^{1/2}$  и  $(p - q) \sim (\Delta t)^{1/2}$ , в силу чего все члены в (24.1), кроме записанных в первой строке, стремятся при  $\Delta t \rightarrow 0$  к нулю. В результате указанный предельный переход приводит к следующему дифференциальному уравнению для плотности вероятности перехода  $v(t, x)$ :

$$\frac{\partial v}{\partial t} = -A \frac{\partial v}{\partial x} + \frac{B}{2} \frac{\partial^2 v}{\partial x^2}. \quad (24.3)$$

Какому начальному условию должно удовлетворять решение при  $t = 0$ ? В общем случае для вероятности перехода  $v(t, x | t_0, x_0)$ , поскольку при  $t = t_0$  с достоверностью должно быть  $x = x_0$ , имеем

$$v(t_0, x | t_0, x_0) = \delta(x - x_0). \quad (24.4)$$

В нашем случае  $v(t, x) \equiv v(t, x | 0, 0)$ , так что надо требовать

$$v(0, x) = \delta(x)$$

и, конечно,

$$\int_{-\infty}^{+\infty} v(t, x) dx = 1.$$

Этими условиями определено решение <sup>1)</sup>

$$v(t, x) = \frac{1}{\sqrt{2\pi Bt}} \exp \left\{ -\frac{(x - At)^2}{2Bt} \right\}, \quad (24.5)$$

т. е. нормальное распределение с  $\bar{x} = At$  и дисперсией  $(x - At)^2 = Bt$ . Для произвольных начальных  $x_0$  и  $t_0$  мы получили бы

$$v(t, x | t_0, x_0) = \frac{\exp \left\{ -\frac{[x - x_0 - A(t - t_0)]^2}{2B(t - t_0)} \right\}}{\sqrt{2\pi B(t - t_0)}}. \quad (24.6)$$

Мы имеем, таким образом, однородный, но нестационарный марковский процесс.

Если бы задача ставилась о блужданиях частицы между двумя *отражающими* стенками, удаленными на расстояние  $l$  друг от друга (при этом, естественно, не должно быть систематического потока, т. е.  $A = 0$ ), то существовала бы предельная стационарная вероятность, а именно равномерное распределение с плотностью  $1/l$ . Мы имели бы

$$\lim_{t \rightarrow \infty} v(t, x | t_0, x_0) = w_1(x) = 1/l.$$

В этом случае налицо и эргодичность процесса, так как  $w_1(x)$  не зависит от начального состояния  $x_0$ . Напротив, при движении частицы между *поглощающими* стенками (прилипание частицы на стенках) предельное распределение существует и стационарно, но эргодичности нет, так как  $w_1(x)$  сохраняет зависимость от  $x_0$ :

$$\lim_{t \rightarrow \infty} v(t, x | t_0, x_0) = w_1(x) = \left(1 - \frac{x_0}{l}\right) \delta(x) + \frac{x_0}{l} \delta(x - l).$$

Нетрудно убедиться в том, что вероятность перехода (24.6) удовлетворяет уравнению Смолуховского (22.5).

Поведение  $v(t, x | t_0, x_0)$  с ростом  $t$  вполне очевидно: распределение постепенно расплывается в соответствии с ростом дисперсии  $B(t - t_0)$ , причем максимум (среднее значение) равномерно перемещается из точки  $x_0$  при  $t = t_0$  с постоянной скоростью  $A$ . Иллюстрацией может служить движение брауновской частицы при наличии силы тяжести. Систематическая скорость равна при этом  $A = -mg/h = \text{const}$ , где  $h$  — коэффициент стокова трения. Таким образом, (24.6) описывает *диффузию в равномерном потоке*, что, впрочем, ясно уже из исходного дифференциального уравнения (24.3). При  $A = 0$  мы возвра-

<sup>1)</sup> Очевидно, что решение — просто теорема Муавра — Лапласа, формулированная применительно к непрерывным «испытаниям» [см. §§ 6, 7, в частности, формулу (7.3)].

щаемся к задаче об *изотропных* блужданиях или же о распределении суммарной *амплитуды*  $x$ , получающейся в результате сложения бесконечно большого числа  $N$  колебаний с *равновероятными* бесконечно малыми амплитудами  $\pm a$ . В последнем случае нас интересует обычно не  $x$ , а интенсивность  $J = x^2$ . Нетрудно подсчитать при помощи формулы (24.5) с  $A = 0$ , что

$$\bar{J} = \overline{x^2} = Bt, \quad \overline{(\Delta J)^2} = \bar{J}^2 - \bar{J}^2 = 2B^2t^2,$$

так что (ср. § 4)

$$\frac{\sqrt{\overline{(\Delta J)^2}}}{\bar{J}} = \sqrt{2}.$$

Относительного сглаживания флуктуаций интенсивности нет: как бы много колебаний мы ни сложили, средний разброс интенсивности будет того же порядка, что и сама средняя интенсивность.

Вопрос об интенсивности, получаемой в результате сложения большого числа колебаний со случайными амплитудами, впервые рассмотрел описанным способом Релей (см. [6], § 42а). Он исследовал также более общий случай равномерного распределения фаз складываемых колебаний в интервале  $(0, 2\pi)$ , т. е. случай *двумерной* векторной диаграммы, или задачу о блужданиях частицы не в одном измерении, а на плоскости (рис. 12). Здесь речь идет о вероятности перехода к моменту  $t$  из начала  $O$  в точку  $r(x, y)$ .

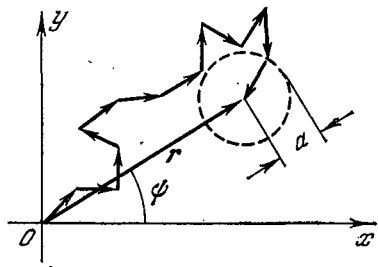


Рис. 12.

Попадание в  $(x, y)$  в момент  $t$  может осуществиться в результате перехода с любой точки окружности радиуса  $a$  (длина шага) с центром в  $(x, y)$ , если частица в момент  $t - \Delta t$  была на этой окружности. При этом все направления элементарного вектора  $\mathbf{a}$  равновероятны (равномерное распределение фазы  $\varphi$ , изотропные блуждания), так что

$$v(t, x, y) = \int_0^{2\pi} v(t - \Delta t, x - a \cos \varphi, y - a \sin \varphi) \frac{d\varphi}{2\pi}.$$

Разлагая подынтегральное выражение по степеням  $a \cos \varphi$  и  $a \sin \varphi$  и выполняя интегрирование, получаем

$$v(t, x, y) = v(t - \Delta t, x, y) + \frac{a^2}{4} \left( \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} \right) + o(a^3).$$

Если предположить теперь, что

$$\lim_{\Delta t \rightarrow 0} \frac{a^2}{\Delta t} = 2B,$$

то в результате предельного перехода  $\Delta t \rightarrow 0$  получаем

$$\frac{\partial v}{\partial t} = \frac{B}{2} \left( \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} \right) \quad (24.7)$$

— двумерное уравнение диффузии, нормированное решение которого, переходящее при  $t = 0$  в  $\delta(x)\delta(y)$ , есть

$$v(t, x, y) = \frac{1}{2\pi Bt} e^{-(x^2+y^2)/2Bt}, \quad (24.8)$$

т. е. произведение нормальных распределений для компонент  $x$  и  $y$ , которые, таким образом, независимы. Наивероятнейшее и среднее значения вектора  $\mathbf{r}$  равны нулю.

Если интересоваться абсолютной величиной вектора  $\mathbf{r}$  и его фазой  $\psi$ , т. е. перейти на плоскости  $(x, y)$  к полярным координатам ( $x = r \cos \psi$ ,  $y = r \sin \psi$ ,  $dx dy = r dr d\psi$ ), то вероятность перехода запишется в виде

$$v(t, x, y) dx dy = \tilde{v}(t, r, \psi) dr d\psi = \frac{e^{-r^2/2Bt}}{Bt} r dr \frac{d\psi}{2\pi}. \quad (24.9)$$

Из (24.9) следует, что  $r$  и  $\psi$  независимы, причем фаза  $\psi$  равномерно распределена в  $(0, 2\pi)$ , а амплитуда  $r$  подчинена закону распределения

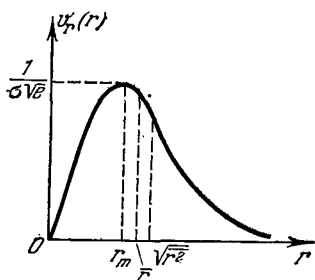


Рис. 13.

$$v_r(t, r) dr = \frac{e^{-r^2/2\sigma^2}}{\sigma^2} r dr, \quad (24.10)$$

называемому *распределением Релея*. Через  $\sigma^2$  мы обозначили среднее квадратичное значение компонент  $x$  и  $y$ :

$$\overline{x^2} = \overline{y^2} = \sigma^2 = Bt.$$

Из (24.10) следует, что наивероятнейшее значение  $r$  есть  $r_m = \sigma$  (рис. 13) и что

$$\bar{r} = \sqrt{\pi/2} \sigma, \quad \bar{r^2} = 2\sigma^2 = 2Bt. \quad (24.11)$$

Последнее выражение вытекает также из того, что  $\bar{r^2} = \overline{x^2} + \overline{y^2}$ .

Изложенный релеевский метод явился первым шагом на пути к установлению общего дифференциального уравнения, которому удовлетворяют вероятности перехода и которое вытекает из уравнения Смолуховского при определенных предположениях об этих вероятностях. Это дифференциальное уравнение



принято называть *уравнением Эйнштейна — Фоккера* по имени ученых, которые впервые его вывели и использовали. А. Н. Колмогоров дал в 1931 г. строгое его обоснование. К выводу и обсуждению этого уравнения мы перейдем в § 26<sup>1)</sup>.

## § 25. Некоторые обобщения распределения Релея

В этом параграфе, хотя он и прерывает изложение теории марковских процессов, мы остановимся на некоторых обобщениях релеевского распределения, к которому нас привела простейшая постановка задачи о случайных блужданиях на плоскости. Отвлекаясь от этой задачи, в которой вектор  $\mathbf{r}$  есть сумма многих случайных шагов  $\mathbf{a}_k$ , можно резюмировать полученный результат следующим образом. Если компоненты  $x$  и  $y$  вектора  $\mathbf{r}$  независимы и распределены нормально с  $\bar{x} = \bar{y} = 0$  и  $\overline{x^2} = \overline{y^2} = \sigma^2$ , то, согласно (24.10), модуль  $r$  и аргумент  $\psi$  этого вектора независимы,  $r$  распределено по Релею, а  $\psi$  — равномерно в интервале  $(0, 2\pi)$ . Отказ от любого из указанных ограничений распределения  $x$  и  $y$  ведет к распределениям  $\mathbf{r}$ , обобщающим релеевское. Рассмотрим сначала случай, когда при сохранении всех остальных предположений допускается, что среднее значение  $\mathbf{r}$  отлично от нуля.

Мы будем говорить о  $\mathbf{r}$  как о комплексной амплитуде колебания на векторной диаграмме ( $r$  — амплитуда,  $\psi$  — фаза колебания).

1. Рассмотрим распределение суммы детерминированного сигнала и нормального шума. Пусть сигнал изображается вектором  $\mathbf{A}$  с постоянными амплитудой  $A$  и фазой  $\psi_0$  (гармонический сигнал), а нормальный шум — вектором  $\mathbf{r}$  с распределением (24.10). Нас интересует распределение вектора  $\mathbf{R} = \mathbf{A} + \mathbf{r}$ , т. е. совместное распределение его амплитуды  $R$  и фазы  $\varphi$  (рис. 14). Так как

$$r \cos \psi = R \cos \varphi - A \cos \psi_0, \quad r \sin \psi = R \sin \varphi - A \sin \psi_0,$$

имеем

$$r^2 = R^2 + A^2 - 2RA \cos(\varphi - \psi_0), \quad r dr d\psi = R dR d\varphi.$$

Поэтому преобразование (24.10) к переменным  $R$  и  $\varphi$  дает

$$w_R(R, \varphi) dR d\varphi =$$

$$= \exp \left\{ -\frac{1}{2\sigma^2} [R^2 + A^2 - 2RA \cos(\varphi - \psi_0)] \right\} \frac{R dR d\varphi}{2\pi\sigma^2}, \quad (25.1)$$

<sup>1)</sup> См. [7], §§ 54—56.

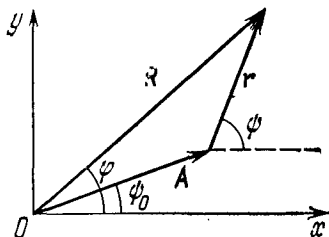


Рис. 14.

что и является искомым распределением. Здесь нет ни независимости  $R$  и  $\varphi$ , ни равномерного распределения  $\varphi$ .

Интегрируя (25.1) по  $\varphi$  от 0 до  $2\pi$  и пользуясь известным выражением для нулевой бесселевой функции мнимого аргумента

$$I_0(a) = J_0(ia) = \frac{1}{2\pi} \int_0^{2\pi} e^{a \cos \alpha} d\alpha, \quad (25.2)$$

получаем распределение амплитуды  $R$ :

$$\omega_R(R) dR = dR \int_0^{2\pi} \omega_R(R, \varphi) d\varphi = I_0\left(\frac{RA}{\sigma^2}\right) e^{-(R^2 + A^2)/2\sigma^2} \frac{R dR}{\sigma^2}, \quad (25.3)$$

которое часто называют распределением Райса [8] или обобщенным релеевским. При  $A = 0$   $\omega_R$  переходит в  $\omega_r$  и, соответственно,  $\omega_R(R)$  — в релеевское распределение. Средний квадрат  $R$  равен  $\bar{R}^2 = A^2 + 2\sigma^2$ , а моменты любого (не обязательно целого) порядка  $\nu$  выражаются формулой

$$\bar{R}^\nu = (2\sigma^2)^{\nu/2} \Gamma\left(1 + \frac{\nu}{2}\right) \cdot {}_1F_1\left(-\frac{\nu}{2}, 1; -\frac{A^2}{2\sigma^2}\right),$$

где  ${}_1F_1$  — вырожденная гипергеометрическая функция. При больших значениях аргумента бесселевой функции, пользуясь асимптотической формулой  $I_0(z) \approx e^z / \sqrt{2\pi z}$  при  $|z| \gg 1$ , получаем из (25.3)

$$\omega_R(R) dR \approx \frac{1}{\sigma} \sqrt{\frac{R}{2\pi A}} e^{-(R-A)^2/2\sigma^2} dR \quad (RA \gg \sigma^2).$$

Интегрируя (25.1) по  $R$  от 0 до  $\infty$ , находим распределение фазы  $\varphi$ :

$$\omega_\varphi(\varphi) d\varphi = e^{-A^2/2\sigma^2} [1 + \sqrt{\pi} \eta e^{\eta^2} (1 + \Phi(\eta))] \frac{d\varphi}{2\pi}, \quad (25.4)$$

где  $\Phi(\eta)$  — так называемый интеграл ошибок:

$$\Phi(\eta) = \frac{2}{\sqrt{\pi}} \int_0^\eta e^{-t^2} dt,$$

и  $\eta = \frac{\sqrt{2} A}{\sigma} \cos(\varphi - \varphi_0)$ . При  $A = 0$  распределение (25.4) становится равномерным.

2. Теперь, сохраняя нормальный закон для  $x$  и  $y$  и предположение, что  $\bar{x} = \bar{y} = 0$ , рассмотрим случай коррелированных  $x$  и  $y$  с неодинаковыми дисперсиями:

$$\overline{x^2} = \sigma_1^2, \quad \overline{y^2} = \sigma_2^2, \quad \overline{xy} = \sigma_1 \sigma_2 K,$$

т. е. нормальное распределение (7.13). Полагая в нем по-прежнему  $x = r \cos \psi$ ,  $y = r \sin \psi$ , получаем

$$\begin{aligned} w_r(r, \psi) dr d\psi = \\ = \exp \left\{ -\frac{r^2}{4(1-K^2)} \left[ \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} + \left( \frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2} \right) \cos 2\psi - \frac{2K}{\sigma_1 \sigma_2} \sin 2\psi \right] \right\} \times \\ \times \frac{r dr d\psi}{2\pi \sigma_1 \sigma_2 \sqrt{1-K^2}}. \quad (25.5) \end{aligned}$$

Интегрирование по  $\psi$  от 0 до  $2\pi$ , если воспользоваться формулой

$$I_0(\sqrt{a^2 + b^2}) = \frac{1}{2\pi} \int_0^{2\pi} e^{(a \cos \alpha + b \sin \alpha)} d\alpha,$$

приводит к распределению

$$w_r(r) dr = I_0 \left[ \frac{qr^2}{4(1-K^2)} \right] \exp \left\{ -\frac{r^2}{4(1-K^2)} \left( \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right) \right\} \frac{r dr}{\sigma_1 \sigma_2 \sqrt{1-K^2}}, \quad (25.6)$$

где

$$q = \sqrt{\left( \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right)^2 - \frac{4(1-K^2)}{\sigma_1^2 \sigma_2^2}}.$$

Если отсутствует корреляция  $x$  и  $y$  ( $K=0$ ), то

$$w_r(r) dr = I_0 \left[ \left( \frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2} \right) \frac{r^2}{4} \right] \exp \left\{ -\left( \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right) \frac{r^2}{4} \right\} \frac{r dr}{\sigma_1 \sigma_2}. \quad (25.7)$$

Если корреляция есть, но  $\sigma_1^2 = \sigma_2^2 = \sigma^2$ , то

$$w_r(r) dr = I_0 \left[ \frac{Kr^2}{2\sigma^2(1-K^2)} \right] \exp \left\{ -\frac{r^2}{2\sigma^2(1-K^2)} \right\} \frac{r dr}{\sigma^2 \sqrt{1-K^2}}. \quad (25.8)$$

Разумеется, обе формулы переходят в релеевское распределение, если в (25.7) положить  $\sigma_1 = \sigma_2$ , а в (25.8) —  $K=0$ .

При коэффициенте корреляции  $K$ , настолько близком к единице, что  $1-K^2 \ll \frac{1}{4} \left( \frac{\sigma_1}{\sigma_2} + \frac{\sigma_2}{\sigma_1} \right)^2$ , и при условии, что аргумент бесселевой функции велик [ $qr^2 \gg 4(1-K^2)$ ], так что можно воспользоваться ее асимптотическим выражением, распределение (25.6) переходит в одностороннее гауссово ( $r \geq 0$ ):

$$w_r(r) dr = 2 \exp \left\{ -\frac{r^2}{2(\sigma_1^2 + \sigma_2^2)} \right\} \frac{dr}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}}.$$

3. Отказ от нормального распределения компонент  $x$  и  $y$ , естественно, уменьшает возможности получения каких-либо конкретных результатов. Рассмотрим случай, когда  $x$  и  $y$

независимы, но распределены по одному и тому же закону, так что их совместное распределение имеет вид  $w(x)dx \cdot w(y)dy$ . Переходя к полярным координатам, получаем

$$w_r(r, \psi) dr d\psi = w(r \cos \psi) w(r \sin \psi) r dr d\psi,$$

откуда

$$w_r(r) dr = r dr \int_0^{2\pi} w(r \cos \psi) w(r \sin \psi) d\psi. \quad (25.9)$$

Конечно, в случае гауссовой плотности  $w$  формула (25.9) возвращает нас к релеевской плотности  $w_r(r)$ .

Введем характеристическую функцию компонент  $x$  и  $y$ :

$$\varphi(u) = \langle e^{iux} \rangle = \int_{-\infty}^{+\infty} e^{iux} w(x) dx.$$

Плотность  $w_r(r)$  можно записать тогда в виде

$$\begin{aligned} w_r(r) &= \frac{r}{4\pi^2} \iint_{-\infty}^{+\infty} \varphi(u) \varphi(v) du dv \int_0^{2\pi} e^{-ir(u \cos \psi + v \sin \psi)} d\psi = \\ &= \frac{r}{2\pi} \iint_{-\infty}^{+\infty} \varphi(u) \varphi(v) J_0(r \sqrt{u^2 + v^2}) du dv. \end{aligned}$$

Если ввести полярные координаты также на плоскости  $(u, v)$ , положив  $u = \lambda \cos \alpha$ ,  $v = \lambda \sin \alpha$  ( $du dv = \lambda d\lambda d\alpha$ ), то последнее выражение преобразуется в следующее:

$$w_r(r) = r \int_0^{\infty} F(\lambda) J_0(\lambda r) \lambda d\lambda, \quad (25.10)$$

где

$$F(\lambda) = \frac{1}{2\pi} \int_0^{2\pi} \varphi(\lambda \cos \alpha) \varphi(\lambda \sin \alpha) d\alpha. \quad (25.11)$$

Таким образом,  $w_r(r)$  представляет собой умноженную на  $r$  трансформанту Бесселя от функции  $F(\lambda)$ .

4. Вернемся теперь к случайным блужданиям на плоскости, когда вектор  $\mathbf{r}$  рассматривается как сумма случайных векторов  $\mathbf{a}_k$  (элементарных скачков или шагов) и статистика  $\mathbf{r}$  не задается непосредственно, а определяется статистикой  $\mathbf{a}_k$ . Возьмем случай изотропных случайных блужданий со случайной длиной шага. Пусть все элементарные векторы  $\mathbf{a}_k$  независимы в совокупности и распределены одинаково, причем полярные углы

(фазы)  $\theta_k$  каждого из них независимы от его амплитуды  $a_k$  и распределены равномерно в интервале  $(0, 2\pi)$  (все направления элементарных шагов равновероятны, что и обеспечивает изотропность блужданий). Таким образом, для любого  $a_k$

$$w_a(a_k, \theta_k) da_k d\theta_k = w_a(a_k) da_k \frac{d\theta_k}{2\pi}.$$

Соответствующая характеристическая функция есть

$$\begin{aligned} \varphi_a(u, v) &= \int_0^\infty w_a(a_k) da_k \cdot \frac{1}{2\pi} \int_0^{2\pi} e^{ia_k(u \cos \theta_k + v \sin \theta_k)} d\theta_k = \\ &= \int_0^\infty w_a(a) J_0(\lambda a) da \equiv \varphi_a(\lambda), \end{aligned} \quad (25.12)$$

где  $\lambda = \sqrt{u^2 + v^2}$ . В силу статистической независимости всех  $a_k$  двумерная характеристическая функция суммарного вектора

$\mathbf{r} = \sum_{k=1}^n \mathbf{a}_k$  равна

$$\varphi_r(u, v) = \varphi_a^n(\lambda) \equiv \varphi_r(\lambda), \quad (25.13)$$

а совместная плотность вероятности декартовых компонент  $x$  и  $y$  вектора  $\mathbf{r}$  запишется в виде

$$\begin{aligned} w_r(x, y) &= \frac{1}{4\pi^2} \iint_{-\infty}^{+\infty} \varphi_r(u, v) e^{-i(ux+vy)} du dv = \\ &= \frac{1}{4\pi^2} \iint_{-\infty}^{+\infty} \varphi_a^n(\lambda) e^{-i(ux+vy)} du dv. \end{aligned}$$

Переходя к полярным координатам как на плоскости  $(x, y)$  ( $x = r \cos \psi$ ,  $y = r \sin \psi$ ), так и на плоскости  $(u, v)$  ( $u = \lambda \cos \alpha$ ,  $v = \lambda \sin \alpha$ ), получаем

$$\begin{aligned} \tilde{w}_r(r, \psi) &= r w_r(r \cos \psi, r \sin \psi) = \\ &= \frac{r}{4\pi^2} \int_0^\infty \varphi_a^n(\lambda) \lambda d\lambda \int_0^{2\pi} e^{-i\lambda r \cos(\psi-\alpha)} d\alpha = \\ &= \frac{r}{2\pi} \int_0^\infty \varphi_a^n(\lambda) J_0(\lambda r) \lambda d\lambda \equiv w_r(r) \frac{1}{2\pi}. \end{aligned} \quad (25.14)$$

Эта совместная плотность  $r$  и  $\psi$  не содержит  $\psi$ , т. е. направление результирующего вектора  $\mathbf{r}$  независимо от его амплитуды  $r$

и распределено равномерно в  $(0, 2\pi)$ . В соответствии с центральной предельной теоремой при  $n \rightarrow \infty$  и  $\overline{r^2} = n\overline{a^2} = 2\sigma^2 = \text{const}$  распределение  $w_r(x, y)$  стремится к нормальному закону для независимых  $x$  и  $y$  с  $\overline{x^2} = \overline{y^2} = \sigma^2$ , а тем самым распределение амплитуды  $\tilde{w}_r(r)$  — к релеевскому [9].

Если вернуться к рассмотренному в § 24 случаю фиксированной длины шага, т. е. к плотности вероятности  $w_a(a_k) = \delta(a_k - a)$ , то по (25.12)  $\Phi_a(\lambda) = J_0(\lambda a)$ . Из (25.14) следует при этом, что

$$w_r(r) = r \int_0^\infty J_0^n(\lambda a) J_0(\lambda r) \lambda d\lambda.$$

Учитывая, что

$$\lambda \int_0^R J_0(\lambda r) r dr = R J_1(\lambda R) = -\frac{dJ_0(\lambda R)}{d\lambda},$$

можно записать вероятность того, что  $r$  не превосходит  $R$ , в виде

$$\begin{aligned} P(r < R) &= \int_0^R w_r(r) dr = R \int_0^\infty J_0^n(\lambda a) J_1(\lambda R) d\lambda = \\ &= - \int_0^\infty J_0^n(\lambda a) \frac{dJ_0(\lambda R)}{d\lambda} d\lambda. \end{aligned}$$

Отсюда при  $R = a$  находим

$$P(r < a) = - \left. \frac{[J_0(\lambda a)]^{n+1}}{n+1} \right|_0^\infty = \frac{1}{n+1}.$$

5. Рассмотрим  $m$ -распределение Накагами [10]. Выше мы рассматривали модуль  $r$  вектора  $\mathbf{r}$ , задавая распределение его декартовых компонент  $x$  и  $y$  ( $r = +\sqrt{x^2 + y^2}$ ), или же полагали, что  $\mathbf{r}$  есть сумма векторов  $\mathbf{a}_k$ , и специальным образом задавали распределение  $\mathbf{a}_k$ . Возьмем теперь существенно более общий случай, когда  $r$  — произвольная неотрицательная детерминированная функция  $2n$ -мерной случайной величины  $\mathbf{a} = \{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ :

$$r = \Phi(\mathbf{a}).$$

Без каких-либо предположений о виде  $2n$ -мерного распределения величины  $\mathbf{a}$  (т. е. совместного распределения декартовых компонент всех  $\mathbf{a}_k$  или распределения всех модулей  $a_k$  и фаз  $\theta_k$ ) имеем (см. задачу 9 гл. II)

$$w_r(r) = \langle \delta[r - \Phi(\mathbf{a})] \rangle, \quad (25.15)$$

где угловые скобки означают усреднение по распределению  $\mathbf{a}$ . При  $\Phi \geq 0$  дельта-функция может быть представлена в виде разложения Ганкеля:

$$\delta(r - \Phi) = r^{\nu+1} \int_0^\infty \frac{J_\nu(\lambda\Phi)}{\Phi^\nu} J_\nu(\lambda r) \lambda d\lambda, \quad \operatorname{Re} \nu \geq -\frac{1}{2}.$$

Подставив это в (25.15), получаем

$$\begin{aligned} w_r(r) &= r^{\nu+1} \int_0^\infty \left\langle \frac{J_\nu(\lambda\Phi)}{\Phi^\nu} \right\rangle J_\nu(\lambda r) \lambda d\lambda = \\ &= \frac{r^{\nu+1}}{2^\nu \Gamma(\nu+1)} \int_0^\infty F_\nu(\lambda) J_\nu(\lambda r) \lambda^{\nu+1} d\lambda, \end{aligned} \quad (25.16)$$

где

$$F_\nu(\lambda) = 2^\nu \Gamma(\nu+1) \left\langle \frac{J_\nu(\lambda\Phi)}{(\lambda\Phi)^\nu} \right\rangle. \quad (25.17)$$

Заметим, что формула (25.10) отвечает частному случаю  $\nu=0$ , когда  $F_0(\lambda) = \langle J_0(\lambda\Phi) \rangle$ .

Из известного степенного разложения

$$\frac{J_\nu(z)}{z^\nu} = \frac{1}{2^\nu} \sum_{k=0}^\infty \frac{(-1)^k (z/2)^{2k}}{k! \Gamma(\nu+k+1)}$$

и из (25.17) следует, что  $F_\nu(\lambda)$  содержит только четные степени  $\lambda$  и только четные моменты  $\Phi$ :

$$F_\nu(\lambda) = \Gamma(\nu+1) \sum_{k=0}^\infty \frac{(-1)^k (\lambda/2)^{2k} \overline{\Phi^{2k}}}{k! \Gamma(\nu+k+1)}. \quad (25.18)$$

При использовании формул (25.16) и (25.18) Накагами считает  $\nu$  вещественным, так что условие, наложенное на  $\operatorname{Re} \nu$ , сводится к требованию  $\nu \geq -1/2$ . Но и при этом ограничении формулы (25.16) и (25.18) охватывают, очевидно, чрезвычайно обширный набор самых разнообразных распределений  $r$ , так как возможен выбор не только значений  $\nu$ , но и распределения многомерной случайной величины  $\mathbf{a}$  — аргумента функции  $\Phi(\mathbf{a})$ . Автор поступает, однако, иначе: он просто *конструирует* функцию  $F_\nu(\lambda)$ , не задаваясь вопросом о том, какому распределению  $\mathbf{a}$  она при этом соответствует. Во-первых, налагается условие, чтобы  $F_\nu(\lambda)$  зависела только от двух параметров, связанных с моментами  $\overline{\Phi^2}$  и  $\overline{\Phi^4}$ , а именно от

$$\Omega = \overline{\Phi^2}, \quad m = \frac{(\overline{\Phi^2})^2}{D[\Phi^2]} = \frac{\Omega^2}{\overline{\Phi^4} - \Omega^2}. \quad (25.19)$$

Во-вторых, известное неравенство Ляпунова

$$\overline{\Phi^{2k+2l}} \geq \overline{\Phi^{2k}} \cdot \overline{\Phi^{2l}}$$

заменяется равенством, с помощью которого высшие четные моменты  $\Phi$  выражаются через  $\overline{\Phi^2}$  и  $\overline{\Phi^4}$  (например,  $\overline{\Phi^6} = \overline{\Phi^2} \cdot \overline{\Phi^4}$ ,  $\overline{\Phi^8} = (\overline{\Phi^4})^2$  и т. д.), т. е. через  $\Omega$  и  $m$ . Тогда (25.18) можно привести к ряду, который в случае  $\nu = 0$  суммируется и дает функцию

$$F_0(\lambda) = e^{-\Omega\lambda^2/4m} P_{m-1}(\Omega\lambda^2/4m), \quad (25.20)$$

где  $P_{m-1}$  — полиномы Лагерра. Если же  $\nu \neq 0$ , то значение  $\nu$  выбирается равным  $m-1$ . Тогда (25.18) приводится к ряду, который приближенно, но с хорошей точностью описывается функцией

$$F_\nu(\lambda) \approx e^{-\Omega\lambda^2/4(\nu+1)}, \quad \nu = m-1. \quad (25.21)$$

Хотя здесь и предполагалось, что  $m \neq 1$ , но, как легко видеть, обе формулы (25.20) и (25.21) при  $m = 1$  ( $\nu = 0$ ) совпадают, поскольку  $P_0 = 1$ .

Подстановка (25.20) (при  $\nu = 0$ ) и (25.21) (при  $\nu = m-1$ ) в (25.16) дает в результате вычисления интегралов одно и то же выражение для  $w_r(r)$ , а именно:

$$w_r(r) \equiv M(r, m, \Omega) = \frac{2m^m r^{2m-1} e^{-mr^2/\Omega}}{\Gamma(m) \Omega^m}, \quad (25.22)$$

что автор и назвал плотностью  $m$ -распределения. Отметим, что в случае  $\nu = m-1$  из условия  $\nu \geq -1/2$  следует, что  $m \geq 1/2$ . Так как по (25.19)  $m = \Omega^2/(r^4 - \Omega^2)$ , где  $\Omega = r^2$ , то это означает, что  $\bar{r}^4 \leq 3(\bar{r}^2)^2$ .

В цитируемой работе [10] рассмотрено много применений и обобщений  $m$ -распределения, но мы ограничимся лишь двумя замечаниями. Из (25.22) видно, что  $m$ -распределение переходит при  $m = 1/2$  в одностороннее гауссово, а при  $m = 1$  — в релеевское распределение. Далее, если перейти в (25.22) к переменной  $x = mr^2/\kappa\Omega$  ( $\kappa > 0$ ) и переобозначить  $m$  через  $\alpha$ , то для  $x$  получается гамма-распределение (10.18)<sup>1)</sup>.

Вернемся теперь к марковским процессам.

## § 26. Непрерывные марковские процессы. Уравнение Эйнштейна — Фоккера

В уравнении Смолуховского (22.1) промежуточный момент времени  $\theta$  может быть выбран между  $t_0$  и  $t$  произвольным образом. Возьмем его весьма близким к  $t$ , положив  $\theta = t - \tau$

<sup>1)</sup> Обобщениям релеевского распределения посвящена работа [11].



(рис. 15), и сделаем предположения о существовании следующих пределов.

Во-первых, мы предполагаем, что

$$\lim_{\tau \rightarrow 0} \frac{\overline{x-y}}{\tau} = \lim_{\tau \rightarrow 0} \frac{1}{\tau} \int_{-\infty}^{+\infty} (x-y) v(t, x|t-\tau, y) dx = A(y, t). \quad (26.1)$$

Смысл этого выражения очевиден:  $\overline{x-y}$  — это условное среднее значение перемещения за время  $\tau$  из фиксированной точки  $y$ , так что  $A(y, t)$  — это средняя скорость изменения состояния в момент  $t$  в точке  $y$  (так называемый коэффициент сноса).

Во-вторых, мы допускаем, что

$$\lim_{\tau \rightarrow 0} \frac{\overline{(x-y)^2}}{\tau} = \lim_{\tau \rightarrow 0} \frac{1}{\tau} \int_{-\infty}^{+\infty} (x-y)^2 \times \\ \times v(t, x|t-\tau, y) dx = B(y, t). \quad (26.2)$$

Величина  $\overline{(x-y)^2}$  есть мера разброса конечных точек  $x$  относительно фиксированной исходной точки  $y$ . Предполагается, таким образом, что этот разброс при удалении от момента  $t-\tau$  на  $\tau$  растет по диффузионному закону, т. е. пропорционально  $\tau$ . Предположения (26.1) и (26.2) воспроизводят условия (24.2), которые мы приняли, когда рассматривали схему с дискретными скачками (§ 24). Коэффициент  $B(y, t)$  или, точнее,  $B/2$  называется коэффициентом диффузии.

В дальнейшем мы подойдем к описанию марковского процесса еще с иной точки зрения, пользуясь представлением о случайных толчках или *случайной силе*, действующей на рассматриваемую систему. Мы увидим тогда, что  $B(y, t)$  характеризует *интенсивность толчков*.

В-третьих, предположим, что

$$\lim_{\tau \rightarrow 0} \frac{\overline{|x-y|^3}}{\tau} = \lim_{\tau \rightarrow 0} \frac{1}{\tau} \int_{-\infty}^{+\infty} |x-y|^3 v(t, x|t-\tau, y) dx = 0. \quad (26.3)$$

Мы считаем, таким образом, что вероятность больших изменений  $|x-y|$  достаточно быстро стремится к нулю при  $\tau \rightarrow 0$ , настолько быстро, что  $\overline{|x-y|^3}$  убывает быстрее  $\tau$ . Именно это требование и позволяет рассматривать  $x$  в системе, подвергающейся действию случайных толчков, как *непрерывно* меняющуюся величину, т. е. как среднее за время, гораздо большее

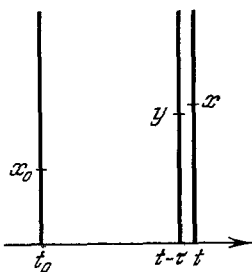


Рис. 15.

промежутка между случайными толчками. Выводы, полученные при предположении (26.3), неприменимы, следовательно, к промежуткам  $\tau$ , малым по сравнению со временем между толчками. Например, вероятность соударения молекул газа за время  $\tau$ , малое по сравнению со временем свободного пробега  $\theta$ , есть  $1 - e^{-\tau/\theta} \approx \tau/\theta$ . При соударении скорость молекулы  $u$  (которая и представляет собой в этом случае величину  $x$ , описывающую состояние молекулы) меняется в среднем на конечную величину  $\Delta u = \pm a$ . Следовательно, для  $\tau \ll \theta$

$$|\overline{\Delta u}|^3 \approx a^3 \tau / \theta$$

и условие

$$\lim_{\tau \rightarrow 0} \frac{|\overline{\Delta u}|^3}{\tau} = 0$$

не выполнено.

Принимая условие (26.3), мы ограничиваемся марковскими процессами, у которых непрерывно не только множество возможных значений, но и само протекание процесса во времени, т. е. смена состояний происходит непрерывно (в вероятностном смысле), без скачков. Такие марковские процессы часто называют *диффузионными*. Случай скачкообразных изменений состояния будет рассмотрен ниже (§ 31).

Умножим уравнение Смолуховского (22.1), положив в нем  $\theta = t - \tau$ , на произвольную функцию  $q(x)$ , обращающуюся вместе с  $q'(x)$  в нуль на границах области (у нас — в  $\pm\infty$ ) и имеющую ограниченную третью производную,  $|q'''(x)| < M$ . Интегрируя по  $x$ , получаем

$$\begin{aligned} \int_{-\infty}^{+\infty} q(x) v(t, x | t_0, x_0) dx &= \\ &= \int_{-\infty}^{+\infty} v(t - \tau, y | t_0, x_0) dy \int_{-\infty}^{+\infty} v(t, x | t - \tau, y) q(x) dx = \\ &= \int_{-\infty}^{+\infty} v(t - \tau, y | t_0, x_0) dy \int_{-\infty}^{+\infty} v(t, x | t - \tau, y) \times \\ &\times \left\{ q(y) + q'(y)(x - y) + q''(y) \frac{(x - y)^2}{2} + q'''(z) \frac{(x - y)^3}{6} \right\} dx, \end{aligned}$$

где  $y < z < x$ . В члене с  $q(y)$  интеграл по  $x$  дает единицу. Перенесем этот член в левую часть, заменим в нем переменную интегрирования  $y$  на  $x$ , разделим все уравнения на  $\tau$  и перей-

дем к пределу при  $\tau \rightarrow 0$ . Это дает

$$\begin{aligned} \int_{-\infty}^{+\infty} q(x) \frac{\partial v(t, x | t_0, x_0)}{\partial t} dx = \\ = \int_{-\infty}^{+\infty} v(t, y | t_0, x_0) \left\{ q'(y) A(y, t) + q''(y) \frac{B(y, t)}{2} + \right. \\ \left. + \frac{q'''(z)}{6} \lim_{\tau \rightarrow 0} \frac{|x-y|^3}{\tau} \right\} dy. \end{aligned}$$

Последний член в правой части ввиду ограниченности  $q'''$  и условия (26.3) равен нулю. Заменяя теперь и справа  $y$  на  $x$  и выполнив интегрирование по частям [с учетом того, что  $q(\pm\infty) = q'(\pm\infty) = 0$ ], получаем ввиду произвольности  $q(x)$

$$\frac{\partial v(t, x | t_0, x_0)}{\partial t} = - \frac{\partial A(x, t) v}{\partial x} + \frac{1}{2} \frac{\partial^2 B(x, t) v}{\partial x^2}. \quad (26.4)$$

Это параболическое уравнение (типа диффузионного) и есть уравнение Эйнштейна — Фоккера (иногда его называют уравнением Фоккера — Планка или вторым уравнением Колмогорова). Полученное ранее уравнение (24.3) представляет собой, очевидно, частный случай (26.4), соответствующий постоянным  $A$  и  $B$ .

Решение уравнения (26.4) должно быть неотрицательно, нормировано к единице и должно удовлетворять начальному условию

$$v(t_0, x | t_0, x_0) = \delta(x - x_0). \quad (26.5)$$

Наглядно уравнение (26.4) можно истолковать следующим образом. В момент  $t_0$  из точки  $x_0$  выходит большое число (ансамбль) частиц, движущихся независимо друг от друга. Их концентрация (относительная «частота») в точке  $x$  в момент  $t$  будет  $v(t, x | t_0, x_0)$ . Поток частиц  $S$  складывается из *систематического* («гидродинамического») потока  $Av$ , где  $A$  — скорость систематического движения в точке  $x$  в момент  $t$ , и из *диффузионного* потока  $-\frac{1}{2} \frac{\partial Bv}{\partial x}$ , где  $B/2$  коэффициент диффузии:

$$S = Av - \frac{1}{2} \frac{\partial Bv}{\partial x}. \quad (26.6)$$

Тогда уравнение Эйнштейна — Фоккера — это просто уравнение

непрерывности:

$$\frac{\partial v}{\partial x} + \frac{\partial S}{\partial x} = 0,$$

выражающее сохранение числа частиц<sup>1)</sup>.

Нетрудно показать, что  $v(t, x | t_0, x_0)$ , рассматриваемая как функция начальных переменных  $x_0, t_0$ , удовлетворяет сопряженному уравнению (в смысле сопряженности линейных операторов  $\hat{L}$  и  $\hat{M}$ , т. е.  $\int_{-\infty}^{+\infty} (f_1 \hat{L} f_2 - f_2 \hat{M} f_1) dx = 0$ ):

$$\frac{\partial v(t, x | t_0, x_0)}{\partial t_0} = -A(x_0, t_0) \frac{\partial v}{\partial x_0} - \frac{B(x_0, t_0)}{2} \frac{\partial^2 v}{\partial x_0^2}. \quad (26.7)$$

Это уравнение называют первым уравнением Колмогорова.

Если в начальный момент  $t_0$  задано не начальное состояние  $x_0$ , а начальное распределение  $w(t_0, x_0)$ , то, поскольку двумерная плотность вероятности есть

$$w_2(t, x; t_0, x_0) = w(t_0, x_0) v(t, x | t_0, x_0),$$

одномерная функция распределения в момент  $t$  будет

$$w_1(t, x) = \int_{-\infty}^{+\infty} w(t_0, x_0) v(t, x | t_0, x_0) dx_0.$$

Умножив (26.4) на  $w(t_0, x_0)$  и проинтегрировав по  $x_0$ , нетрудно убедиться, что вероятность состояния  $w_1(t, x)$  удовлетворяет тому же уравнению:

$$\frac{\partial w_1(t, x)}{\partial t} = -\frac{\partial A(x, t) w_1}{\partial x} + \frac{1}{2} \frac{\partial^2 B(x, t) w_1}{\partial x^2}. \quad (26.8)$$

Начальным условием здесь, конечно, будет

$$w_1(x, t_0) = w(x, t_0), \quad (26.9)$$

причем решение также должно быть неотрицательно и нормировано к единице. Рассмотрим некоторые следствия и частные случаи уравнения (26.4).

<sup>1)</sup> Обычно коэффициентом диффузии называется множитель при  $-\partial v / \partial x$ . Поэтому в (26.6), строго говоря, нельзя назвать  $B/2$  коэффициентом диффузии. Если же записать полный поток в виде

$$S = \left( A - \frac{1}{2} \frac{\partial B}{\partial x} \right) v - \frac{B}{2} \frac{\partial v}{\partial x},$$

то  $B/2$  — коэффициент диффузии, но тогда систематическая часть  $S$  содержит градиент  $B$ .

Введем условные моменты

$$\overline{x^n(t)} = \int_{-\infty}^{+\infty} x^n v(t, x | t_0, x_0) dx,$$

т. е. средние значения  $x^n$  в момент  $t$  при условии, что  $\bar{x}(t_0) = x_0$ . Умножим (26.4) на  $x^n$  и проинтегрируем по  $x$  от  $-\infty$  до  $+\infty$ . Интегрируя члены с  $A$  и  $B$  по частям и предполагая, что проинтегрированные выражения исчезают на границах (у нас — в  $\pm \infty$ ), получаем

$$\frac{d\overline{x^n}}{dt} = n\overline{Ax^{n-1}} + \frac{n(n-1)}{2} \overline{Bx^{n-2}}, \quad n = 1, 2, \dots \quad (26.10)$$

В частности, для  $n = 1$  и 2

$$\frac{d\bar{x}}{dt} = \overline{A(x, t)}, \quad (26.11)$$

$$\frac{d\overline{x^2}}{dt} = 2\overline{A(x, t)x} + \overline{B(x, t)}. \quad (26.12)$$

Подчеркнем для ясности, что функции  $A(x, t)$  и  $B(x, t)$ , согласно их определениям (26.1) и (26.2), — это *детерминированные* функции. Черта над ними в (26.10) — (26.12) означает усреднение по условному распределению  $v(t, x | t_0, x_0) dx$  случайного аргумента  $x$  этих функций.

Если  $A$  и  $B$  — ряды (или полиномы) по степеням  $x$ , то (26.10) представляет собой систему обыкновенных линейных дифференциальных уравнений для моментов  $\overline{x^n}$ . В общем случае все уравнения этой бесконечной системы «зацепляются», т. е. содержат моменты разных порядков. Это относится и к уравнению (26.10). Хотя из него и выпало влияние диффузионного потока вероятности —  $\frac{1}{2} \frac{\partial Bv}{\partial x}$ , оно не является уравнением феноменологической кинетики или динамики, которое должно быть замкнутым (содержащим только момент  $\bar{x}$ ). Исключением является только тот случай, когда  $A(x, t)$  — линейная функция  $x$ :  $A(x, t) = a(t)x + b(t)$ . Тогда уравнение (26.11) отщепляется и становится замкнутым:

$$\frac{d\bar{x}}{dt} = a(t) \bar{x} + b(t).$$

Во всяком случае из (26.11) мы можем сделать вывод, что  $dx/dt$  и  $A(x, t)$  могут различаться только слагаемым, среднее значение которого равно нулю:

$$\frac{dx}{dt} = A(x, t) + F(x, t), \quad \overline{F(x, t)} = 0, \quad (26.13)$$

— замечание, которым мы еще воспользуемся (§ 37).

Возьмем теперь *стационарный* марковский процесс. Вероятность перехода зависит в этом случае от  $t - t_0$ , так что, согласно (26.1) и (26.2),  $A$  и  $B$  не зависят от  $t$ . Одномерная вероятность состояния, *если она существует*, вообще не зависит от  $t$  (§ 16). В результате для  $w_1(x)$  получаем из (26.8) уравнение

$$\frac{d}{dx} \left( Aw_1 - \frac{1}{2} \frac{dBw_1}{dx} \right) = 0.$$

Если на границах области изменения  $x$  стационарный поток

$$S = Aw_1 - \frac{1}{2} \frac{dBw_1}{dx} = 0,$$

то интегрирование дает, что он равен нулю всюду:

$$Aw_1 - \frac{1}{2} \frac{dBw_1}{dx} = 0.$$

Для бесконечной области изменения  $x$  условия  $w = w' = 0$  при  $|x| = \infty$  имеют место при положительном и отличном от нуля  $B$  и при  $A > 0$  или  $A < 0$  соответственно для  $x \rightarrow -\infty$  или  $x \rightarrow +\infty$ , причем в обоих случаях  $|A| > C > 0$  [19]<sup>1)</sup>. Интегрируя последнее уравнение, получаем

$$w_1(x) = \frac{C}{B(x)} \exp \left\{ \int_0^x \frac{2A(s)}{B(s)} ds \right\}. \quad (26.14)$$

Постоянная  $C$  определяется из условия нормировки.

Примером, когда стационарное распределение существует, может служить броуновское движение частицы при наличии силы тяжести ( $A = -mg/h = \text{const}$ ) над *отражающей границей*. Очевидно, на отражающей границе должно быть выполнено именно условие обращения потока  $S$  в нуль, так что выражение (26.14) справедливо; при постоянных  $A$  и  $B$  оно дает

$$w_1(x) = C \exp \left\{ -\frac{2mgx}{hB} \right\},$$

т. е. барометрическую формулу. Как известно (см. [7], § 53),

$$B = 2kT/h,$$

где  $k$  — постоянная Больцмана,  $T$  — абсолютная температура. Следовательно,

$$w_1(x) = C \exp \left\{ -\frac{mgx}{kT} \right\}.$$

Стационарная вероятность  $w_1(x)$  соответствует именно той вероятности состояния, которая рассматривается в статистиче-

<sup>1)</sup> Если эти условия не выполнены, то стационарного решения может не существовать [см. формулу (26.14)].

ской термодинамике. Для изотермической системы распределение  $\omega_1$  является каноническим, для адиабатически изолированной системы — микроканоническим.

Укажем еще на *процесс, однородный по абсциссе* (процесс типа Башелье). В этом случае вероятность перехода зависит только от  $x - x_0$  и, следовательно, по (26.1) и (26.2),  $A$  и  $B$  могут зависеть от  $t$ , но не от  $x$ . Уравнение (26.4) принимает вид

$$\frac{\partial v(t, x - x_0, t_0)}{\partial t} = -A(t) \frac{\partial v}{\partial x} + \frac{B(t)}{2} \frac{\partial^2 v}{\partial x^2}. \quad (26.15)$$

Замена переменных

$$x' = x - \int_{t_0}^t A(\theta) d\theta, \quad t' = \int_{t_0}^t B(\theta) d\theta$$

приводит это уравнение к виду

$$\frac{\partial v}{\partial t'} = \frac{1}{2} \frac{\partial^2 v}{\partial x'^2},$$

так что в первоначальных переменных решение для неограниченной области есть

$$v = \frac{1}{\sqrt{2\pi\sigma(t)}} \exp \left\{ -\frac{1}{2\sigma^2(t)} \left[ x - x_0 - \int_{t_0}^t A(\theta) d\theta \right]^2 \right\},$$

$$\sigma^2(t) = \int_{t_0}^t B(\theta) d\theta.$$

В работе [12] подробно исследована гораздо более общая постановка вопроса — о дифференциальном уравнении для плотности условной вероятности  $v(t, x|T, X)$  любого случайного процесса с непрерывным множеством состояний. Под  $(T, X)$  понимается совокупность условий

$$(T, X) = (t_1, x_1; \dots; t_m, x_m; t_{m+1}, y_{m+1}; \dots; t_k, z_k),$$

причем  $x(t), y(t), \dots, z(t)$  могут быть и *разными* случайными процессами, а множество  $T$  моментов времени  $t_1, \dots, t_k$  может быть и *неупорядоченным*. Существенно только то, что момент времени  $t$  не принадлежит к множеству  $T$ .

Исходным является, конечно, уже не уравнение Смолуховского, а формула для полной плотности условных вероятностей:

$$v(t, x|T, X) = \int_{-\infty}^{+\infty} \bar{v}(t, x|t-\tau, y; T, X) v(t-\tau, y|T, X) dy, \quad (26.16)$$

где  $\tilde{v}$  и  $v$ , вообще говоря, различные функции. Показано (см. задачу 16), что для плотности условной вероятности  $v(t, x|T, X)$  справедливо уравнение вида

$$\frac{\partial v(t, x|T, X)}{\partial t} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n A_n v}{\partial x^n}, \quad (26.17)$$

$$A_n \equiv A_n(x, t|T, X) = \lim_{\tau \rightarrow 0} \frac{1}{\tau} \int (x-y)^n \tilde{v}(t, x|t-\tau, y; T, X) dy. \quad (26.18)$$

Допустимо, что пределы справа и слева ( $\tau \rightarrow 0 \pm$ ) могут быть здесь различны, но, по предположению, они конечны. Для коэффициентов  $A_n$  показано, что если для какого-нибудь четного  $n$  коэффициент  $A_n = 0$ , то  $A_n = 0$  для всех  $n \geq 3$ . Другими словами, если порядок уравнения (26.17) конечен, то он должен быть не выше второго, и тогда получается обобщенное уравнение Эйнштейна — Фоккера:

$$\frac{\partial v(t, x|T, X)}{\partial t} = - \frac{\partial A_1 v}{\partial x} + \frac{1}{2} \frac{\partial^2 A_2 v}{\partial x^2}. \quad (26.19)$$

В частном случае марковских процессов совокупность условий  $(T, X)$  сводится к какому-либо одному предшествующему (или последующему) состоянию самого процесса  $x(t)$ , а функция  $\tilde{v}$  совпадает с  $v$ . Указанная теорема об обращении в нуль всех  $A_n$  с  $n \geq 3$ , разумеется, остается в силе, так что уравнение (26.19) превращается в диффузионное уравнение (26.4) для плотности вероятности перехода  $v(t, x|t_0, x_0)$ .

В принципе, если известны коэффициенты  $A_n$  вида

$$A_n(x_k, t_k|t_1, x_1; \dots; t_{k-1}, x_{k-1}), \quad k = 1, 2, \dots,$$

уравнения (26.17) или (26.19) позволяют найти все условные плотности  $v_k(t_k, x_k|t_1, x_1; \dots; t_{k-1}, x_{k-1})$ , а значит, и любую  $n$ -мерную функцию распределения:

$$\omega_n(t_1, x_1; \dots; t_n, x_n) = \omega_1(t_1, x_1) \prod_{k=2}^n v_k(t_k, x_k|t_1, x_1; \dots; t_{k-1}, x_{k-1}).$$

Предполагается, что  $\omega_1$  существует и либо известно, либо может быть получено из условной плотности вероятностей  $v_k$  при отодвигании всех или части времен  $t_1, \dots, t_{k-1}$  в  $-\infty$  (выполнение сильного условия эргодичности). Однако реализация этой программы нахождения  $\omega_n$  связана с фундаментальной трудностью, на которую наталкиваются все рассматриваемые чисто вероятностные схемы, в том числе и классическая схема для марковских процессов. Ведь регулярная процедура нахождения коэффициентов  $A_n$  требует, чтобы мы уже располагали искомой



условной плотностью вероятностей  $\bar{v}$  (хотя бы для сколь угодно малых времен перехода  $\tau$ ). Сверх того, если допустить, что  $A_n$  откуда-то известны, имеется немало и других трудностей, но они носят более технический характер. Они связаны со сложностью самих уравнений, с необходимостью разбивать область возможных значений  $x(t)$  на участки, в которых  $v$  и ее производные непрерывны, с установлением правильных граничных условий для искомого решения и т. д.

К тому, что уравнения вида (26.17) или (26.19) верны и для немарковских процессов, независимо пришел ряд авторов, наряду с автором цитируемой работы [12], в которой проведен наиболее полный анализ вопроса. Теория легко распространяется и на многомерные случайные процессы. Переходя в следующем параграфе к этому обобщению, мы вернемся к марковским процессам и снова ограничимся диффузионным уравнением Эйнштейна — Фоккера.

### § 27. Обобщение на многомерные случайные функции

Пусть состояние системы в момент  $t$  описывается совокупностью  $n$  случайных величин  $x_i(t)$ ,  $i = 1, 2, \dots, n$ , т. е. случайным «вектором»  $\mathbf{x}(t)$  в  $n$ -мерном пространстве. Компоненты случайной функции  $\mathbf{x}(t)$  могут представлять собой (обобщенные) координаты системы либо же совокупность координат и скоростей (или импульсов). «Координаты»  $x_i$  в  $n$ -мерном пространстве, вообще говоря, криволинейны, но вывод уравнения Эйнштейна — Фоккера не зависит от выбора координат и вообще от метрики пространства  $\mathbf{x}$ , поскольку исходное уравнение Смолуховского

$$v(t, \mathbf{x} | t_0, \mathbf{x}_0) d\mathbf{x} = d\mathbf{x} \int v(t, \mathbf{x} | t - \tau, \mathbf{y}) v(t - \tau, \mathbf{y} | t_0, \mathbf{x}_0) d\mathbf{y}$$

связывает вероятности  $dP = v d\mathbf{x}$ , от которых естественно требовать, чтобы они были скалярами при любом преобразовании координат. Все рассуждения остаются теми же, что и в предыдущем параграфе, с тем только отличием, что теперь речь идет о вероятностях перехода для  $n$ -мерных случайных величин.

Аналогично одномерному случаю делаются предположения

$$\begin{aligned} \lim_{\tau \rightarrow 0} \frac{x_i - y_i}{\tau} &= \lim_{\tau \rightarrow 0} \frac{1}{\tau} \int (x_i - y_i) v(t, z | t - \tau, \mathbf{y}) d\mathbf{x} = A_i(\mathbf{y}, t), \\ \lim_{\tau \rightarrow 0} \frac{(x_i - y_i)(x_k - y_k)}{\tau} &= B_{ik}(\mathbf{y}, t), \\ \lim_{\tau \rightarrow 0} \frac{1}{\tau} |(x_i - y_i)(x_k - y_k)(x_l - y_l)| &= 0, \end{aligned} \quad (27.1)$$

что позволяет прежним путем вывести уравнение Эйнштейна — Фоккера:

$$\frac{\partial v}{\partial t} = - \sum_i \frac{\partial A_i v}{\partial x_i} + \frac{1}{2} \sum_{i, k} \frac{\partial^2 B_{ik} v}{\partial x_i \partial x_k}. \quad (27.2)$$

Решение этого уравнения должно быть неотрицательно, нормировано к единице:

$$\int v(t, \mathbf{x} | t_0, \mathbf{x}_0) d\mathbf{x} = 1, \quad (27.3)$$

и должно удовлетворять начальному условию:

$$\begin{aligned} v(t_0, \mathbf{x} | t_0, \mathbf{x}_0) &= \delta(\mathbf{x} - \mathbf{x}_0) = \\ &= \delta(x_1 - x_{10}) \delta(x_2 - x_{20}) \dots \delta(x_n - x_{n0}). \end{aligned} \quad (27.4)$$

Кроме того, могут быть наложены те или иные граничные условия.

Как уже было отмечено, вид уравнения (27.2) одинаков при любом выборе «координат»  $\mathbf{x}$ , т. е. должен сохраняться при всяком преобразовании  $x_i = x_i(\mathbf{x}')$ , удовлетворяющем, конечно, условиям взаимной однозначности и непрерывности. Однако ковариантность уравнения в целом не означает, что инвариантны величины, через которые оно записано. Действительно, из того, что вероятность  $dP$  — скаляр, вытекает неинвариантность плотности вероятности  $v$ , поскольку инвариантный элемент объема равен не  $d\mathbf{x} = dx_1 \dots dx_n$ , а  $\sqrt{g} d\mathbf{x}$ , где  $g$  — детерминант, составленный из элементов метрического тензора  $g_{ik}$ . Можно определить скалярную плотность вероятности  $\tilde{v} = v / \sqrt{g}$ :

$$dP = v d\mathbf{x} = \tilde{v} \sqrt{g} d\mathbf{x},$$

которой и удобно пользоваться при преобразованиях координат. В частности, если пространство  $\mathbf{x}$  —  $2n$ -мерное фазовое пространство ( $n$  координат  $x_1, \dots, x_n$  и  $n$  скоростей  $\dot{x}_1, \dots, \dot{x}_n$  или импульсов  $p_1, \dots, p_n$ ), то  $\tilde{v}$  представляет собой *фазовую* плотность вероятности.

Далее, при переходе к новым координатам  $x'_\alpha$  коэффициенты  $A_i$  и  $B_{ik}$ , как это нетрудно установить при помощи формул (27.1), преобразуются следующим образом:

$$\begin{aligned} A_i &= \sum_\alpha A'_\alpha \frac{\partial x_i}{\partial x'_\alpha} + \frac{1}{2} \sum_{\alpha, \beta} B'_{\alpha\beta} \frac{\partial^2 x_i}{\partial x'_\alpha \partial x'_\beta}, \\ B_{ik} &= \sum_{\alpha, \beta} B'_{\alpha\beta} \frac{\partial x_i}{\partial x'_\alpha} \frac{\partial x_k}{\partial x'_\beta}, \end{aligned} \quad (27.5)$$

откуда видно, что  $B_{ik}$  — тензор второго ранга, но  $A_i$  не образуют вектора (если не ограничиваться линейными преобразованиями). Это значит, что равенство величин  $A_i$  нулю в одной системе координат не означает их равенства нулю в другой, т. е.  $A_i$  не характеризуют скорость систематического потока инвариантным образом. Однако, в согласии со сказанным выше, при правильных законах преобразования всех величин (включая, конечно, и операции дифференцирования по  $x_i$ ) можно убедиться в ковариантности уравнения (27.2) и непосредственно, т. е. не апеллируя к его выводу<sup>1)</sup>.

Запишем уравнение (27.2) для следующих двух случаев, которые нам вскоре понадобятся, — для частицы, испытывающей действие однородных и изотропных толчков, движущейся 1) на плоскости и 2) по поверхности единичной сферы.

В первом случае (на плоскости) нас будет интересовать уравнение (27.2), записанное в полярных координатах  $r$  и  $\varphi$ :

$$x_1 \equiv x = r \cos \varphi, \quad x_2 \equiv y = r \sin \varphi.$$

Изотропность толчков означает, что средний квадрат смещения одинаков по любому направлению:

$$\overline{(\Delta x)^2} = \overline{(\Delta y)^2} = \overline{(\Delta r)^2} = r^2 \overline{(\Delta \varphi)^2} = \varepsilon^2, \quad (27.6)$$

что возможно только при условии некоррелированности взаимно ортогональных смещений:

$$\overline{\Delta x \cdot \Delta y} = \overline{\Delta r \cdot r \Delta \varphi} = 0. \quad (27.7)$$

Однородность толчков означает, что средний квадрат смещения один и тот же в любом месте плоскости, т. е.  $\varepsilon^2$  не зависит от координат  $x$ ,  $y$  или  $r$ ,  $\varphi$ .

Согласно (27.2) имеем

$$\frac{\partial v}{\partial t} = -\frac{\partial A_r v}{\partial r} - \frac{\partial A_\varphi v}{\partial \varphi} + \frac{1}{2} \left( \frac{\partial^2 B_{rr} v}{\partial r^2} + 2 \frac{\partial^2 B_{r\varphi} v}{\partial r \partial \varphi} + \frac{\partial^2 B_{\varphi\varphi} v}{\partial \varphi^2} \right), \quad (27.8)$$

где, в соответствии с (27.1),

$$A_r = \lim_{\Delta t \rightarrow 0} \frac{\overline{\Delta r}}{\Delta t}, \quad A_\varphi = \lim_{\Delta t \rightarrow 0} \frac{\overline{\Delta \varphi}}{\Delta t}, \quad B_{rr} = \lim_{\Delta t \rightarrow 0} \frac{\overline{(\Delta r)^2}}{\Delta t},$$

$$B_{r\varphi} = \lim_{\Delta t \rightarrow 0} \frac{\overline{\Delta r \Delta \varphi}}{\Delta t}, \quad B_{\varphi\varphi} = \lim_{\Delta t \rightarrow 0} \frac{\overline{(\Delta \varphi)^2}}{\Delta t}.$$

<sup>1)</sup> При произвольных преобразованиях координат необходимо, конечно, различать ко- и контравариантные компоненты. К вопросу о записи уравнения Эйнштейна — Фоккера через инвариантные характеристики марковского процесса см. [13] и [14]. Заметим, что вектором является величина  $\tilde{A}^i = A^i + \frac{1}{2} \Gamma_{rs}^i B^{rs}$ , где  $\Gamma_{rs}^i$  — символы Кристоффеля. См. также [15], § 2. 3.

Из (27.6) и (27.7) следует, что

$$B_{rr} = r^2 B_{\varphi\varphi} = \lim_{\Delta t \rightarrow 0} \frac{\varepsilon^2}{\Delta t} = B, \quad B_{r\varphi} = 0. \quad (27.9)$$

Что касается  $A_r$  и  $A_\varphi$ , то их вычисление следует проводить более аккуратно, не ограничиваясь членами первого порядка относительно  $\varepsilon$ , а учитывая, в соответствии с (25.7), и второй. Из формул

$$(r + \Delta r)^2 = (x + \Delta x)^2 + (y + \Delta y)^2, \quad \varphi + \Delta\varphi = \arctg \frac{y + \Delta y}{x + \Delta x}$$

имеем

$$\begin{aligned} \Delta r &= \Delta x \cos \varphi + \Delta y \sin \varphi + \\ &\quad + \frac{(\Delta x)^2 \sin^2 \varphi + (\Delta y)^2 \cos^2 \varphi - 2 \Delta x \Delta y \sin \varphi \cos \varphi}{2r} + \dots, \\ r \Delta \varphi &= \Delta y \cos \varphi - \Delta x \sin \varphi + \\ &\quad + \frac{[(\Delta x)^2 - (\Delta y)^2] \sin \varphi \cos \varphi - \Delta x \Delta y (\cos^2 \varphi - \sin^2 \varphi)}{r} + \dots \end{aligned}$$

Усредняя эти равенства, деля их на  $\Delta t$  и переходя к пределу при  $\Delta t \rightarrow 0$ , получаем с учетом (27.6) и (27.7)

$$A_r = R + \frac{B}{2r}, \quad r A_\varphi = \Phi, \quad (27.10)$$

где введены величины

$$\begin{aligned} R &= A_x \cos \varphi + A_y \sin \varphi, \quad \Phi = A_y \cos \varphi - A_x \sin \varphi, \\ A_x &= \lim_{\Delta t \rightarrow 0} \frac{\Delta x}{\Delta t}, \quad A_y = \lim_{\Delta t \rightarrow 0} \frac{\Delta y}{\Delta t}. \end{aligned} \quad (27.11)$$

Если бы  $A_x$  и  $A_y$  были декартовыми компонентами вектора  $\mathbf{A}$ , то его полярные компоненты  $A_r$  и  $r A_\varphi$  совпадали бы соответственно с  $R$  и  $\Phi$ . Однако первое равенство (27.10) показывает, что  $A_r$  не совпадает с  $r$ -компонентой вектора  $\mathbf{A}$ . Легко видеть, что подстановка (27.9) и (27.10) в (27.8) приводит снова к уравнению

$$\frac{\partial v}{\partial t} = -\frac{\partial R v}{\partial r} - \frac{\partial \Phi v}{r \partial \varphi} + \frac{B}{2} \left[ \frac{\partial}{\partial r} \left( r \frac{\partial}{\partial r} \left( \frac{v}{r} \right) \right) + \frac{1}{r^2} \frac{\partial^2 v}{\partial \varphi^2} \right], \quad (27.12)$$

где  $v = v(t, r, \varphi | t_0, r_0, \varphi_0)$  — обычная (неинвариантная) плотность вероятности перехода, т. е.  $dP = v \cdot dr d\varphi$ . Заметим, что для скалярной плотности  $\tilde{v} = v/r$ , для которой  $dP = \tilde{v} \cdot r dr d\varphi$ , уравнение (27.12) принимает вид

$$\begin{aligned} \frac{\partial \tilde{v}}{\partial t} &= -\frac{1}{r} \frac{\partial R r \tilde{v}}{\partial r} - \frac{\partial \Phi \tilde{v}}{r \partial \varphi} + \frac{B}{2} \left[ \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial \tilde{v}}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 \tilde{v}}{\partial \varphi^2} \right] = \\ &= -\operatorname{div}(\mathbf{A} \tilde{v}) + \frac{B}{2} \Delta \tilde{v}, \end{aligned}$$

где под  $\mathbf{A}$  понимается вектор с полярными компонентами  $R$  и  $\Phi$ .

Обратимся теперь к движению частицы по единичной сфере. Координатами частицы будут полярный угол  $\theta$  и азимут  $\varphi$ , так что уравнение (27.2) будет

$$\frac{\partial v(t, \theta, \varphi | t_0, \theta_0, \varphi_0)}{\partial t} = -\frac{\partial A_\theta v}{\partial \theta} - \frac{\partial A_\varphi v}{\partial \varphi} + \frac{1}{2} \left( \frac{\partial^2 B_{\theta\theta} v}{\partial \theta^2} + 2 \frac{\partial^2 B_{\theta\varphi} v}{\partial \theta \partial \varphi} + \frac{\partial^2 B_{\varphi\varphi} v}{\partial \varphi^2} \right), \quad (27.13)$$

где

$$A_\theta = \lim_{\Delta t \rightarrow 0} \frac{\overline{\Delta \theta}}{\Delta t}, \quad A_\varphi = \lim_{\Delta t \rightarrow 0} \frac{\overline{\Delta \varphi}}{\Delta t}, \quad B_{\theta\theta} = \lim_{\Delta t \rightarrow 0} \frac{(\overline{\Delta \theta})^2}{\Delta t},$$

$$B_{\theta\varphi} = \lim_{\Delta t \rightarrow 0} \frac{\overline{\Delta \theta \Delta \varphi}}{\Delta t}, \quad B_{\varphi\varphi} = \lim_{\Delta t \rightarrow 0} \frac{(\overline{\Delta \varphi})^2}{\Delta t}.$$

Это уравнение определяет обычную плотность вероятности перехода  $v(t, \theta, \varphi | t_0, \theta_0, \varphi_0)$ , т. е.  $dP = v d\theta d\varphi$ .

В силу предположенной изотропности толчков имеем

$$(\overline{\Delta \theta})^2 = \sin^2 \theta (\overline{\Delta \varphi})^2 = \varepsilon^2, \quad \sin \theta \overline{\Delta \varphi \Delta \theta} = 0,$$

так что

$$B_{\theta\theta} = \sin^2 \theta B_{\varphi\varphi} = \lim_{\Delta t \rightarrow 0} \frac{\varepsilon^2}{\Delta t} \equiv B, \quad B_{\theta\varphi} = 0,$$

причем  $B$  на всей сфере постоянно (однородность толчков).

Если частица совершает малое перемещение  $\Delta\alpha$  из точки  $(\theta, \varphi)$  в точку  $(\theta', \varphi')$ , то по формуле сферической тригонометрии имеем

$$\cos \theta' = \cos \theta \cos \Delta\alpha + \sin \theta \cos \psi \sin \Delta\alpha,$$

где  $\psi$  — угол поворота дуги  $\Delta\alpha$ , отсчитанный от меридиана (рис. 16). Разлагая  $\cos \theta'$  по степеням  $\Delta\theta = \theta' - \theta$  и ограничиваясь вторым порядком относительно  $\Delta\alpha$ , получаем

$$\Delta\theta = \operatorname{ctg} \theta \sin^2 \psi \frac{(\Delta\alpha)^2}{2} - \cos \psi \Delta\alpha.$$

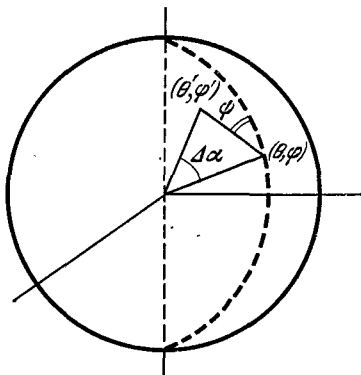


Рис. 16.

Очевидно, все направления  $\psi$  от 0 до  $2\pi$  равновероятны (изотропность толчков), так что  $\cos \psi = 0$ ,  $\sin^2 \psi = 1/2$ . С другой стороны, рассматривая малый участок сферы около точки  $(\theta, \varphi)$ , как плоскость, имеем

$$(\Delta\alpha)^2 = (\Delta\theta)^2 + \sin^2 \theta (\Delta\varphi)^2,$$

откуда

$$\overline{(\Delta\alpha)^2} = \overline{(\Delta\theta)^2} + \sin^2 \theta \overline{(\Delta\varphi)^2} = 2\varepsilon^2.$$

В результате

$$\overline{\Delta\theta} = \operatorname{ctg} \theta \frac{\varepsilon^2}{2}, \quad \text{т. е.} \quad A_\theta = \frac{B}{2} \operatorname{ctg} \theta.$$

Аналогичным путем нетрудно установить, что  $A_\varphi = 0$ . Уравнение (27.13), в котором

$$\begin{aligned} A_\theta &= \frac{B}{2} \operatorname{ctg} \theta, & A_\varphi &= 0, & B_{\theta\theta} &= B, \\ B_{\varphi\varphi} &= \frac{B}{\sin^2 \theta}, & B_{\theta\varphi} &= 0, \end{aligned} \quad (27.14)$$

принимает вид

$$\frac{\partial v}{\partial t} = \frac{B}{2} \left[ \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \left( \frac{v}{\sin \theta} \right) \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 v}{\partial \varphi^2} \right], \quad (27.15)$$

а для скалярной плотности  $\tilde{v}$ , т. е. плотности в элементе телесного угла  $\sin \theta d\theta d\varphi$  ( $\tilde{v} = v/\sin \theta$ ), уравнение будет

$$\sin \theta \frac{\partial \tilde{v}}{\partial t} = \frac{B}{2} \left[ \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial \tilde{v}}{\partial \theta} \right) + \frac{1}{\sin \theta} \frac{\partial^2 \tilde{v}}{\partial \varphi^2} \right]. \quad (27.16)$$

Если отсчитывать  $\theta$  от начального положения частицы, т. е. провести полярную ось через это положение (таким образом  $\theta_0 = 0$ ), то в силу симметрии  $\tilde{v}$  будет независима от  $\varphi$ . Уравнение примет тогда вид

$$\sin \theta \frac{\partial \tilde{v}(t, \theta | t_0, 0)}{\partial t} = \frac{B}{2} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial \tilde{v}}{\partial \theta} \right), \quad (27.17)$$

причем решение надо искать при начальном условии

$$\tilde{v}_{t=t_0} = \frac{\delta(\theta)}{2\pi \sin \theta}.$$

Решение выражается в полиномах Лежандра (§ 30).

## § 28. Флуктуации в томсоновском ламповом генераторе

Рассмотрим применение уравнения Эйнштейна — Фоккера к простейшей автоколебательной системе с одной степенью свободы — ламповому генератору с колебательным контуром в анодной цепи, индуктивной обратной связью и кубической характеристикой лампы (мягкий режим самовозбуждения). Схема генератора показана на рис. 17. Говоря, что генератор является *томсоновской* системой, имеют в виду его близость к линейному гармоническому осциллятору, т. е. достаточную малость диссипативных и нелинейных членов в описывающем генератор дифференциальном уравнении. Конкретные предположения о схеме генератора имеют целью только упрощение динамической мо-

дели, излагаемая же теория не связана рамками именно такой схемы, а применима в общем случае томсоновской автоколебательной системы, как автономной, так и неавтономной, и не только с одной степенью свободы, когда динамическое уравнение имеет вид

$$\ddot{x} + x = \mu f(x, \dot{x}, t)$$

( $\mu$  — малый параметр), но и со многими степенями свободы.

Равным образом не играют роли происхождения и конкретный вид нелинейности, т. е. теория приложима и к другим ламповым схемам, и к генераторам на полупроводниковых элементах, и к оптическим квантовым генераторам (в тех случаях, когда допустимо их квазиклассическое описание), и т. д.

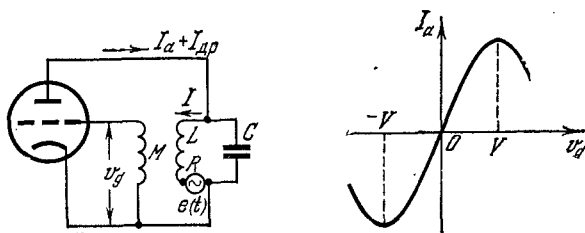


Рис. 17.

Имея в виду последующее использование рассматриваемой модели (§ 53), мы учтем при составлении дифференциального уравнения генератора два источника флуктуаций — *дробовой ток* в анодной цепи лампы  $I_{др}(t)$ , добавляющийся к мгновенному среднему (т. е. статистическому среднему, взятому по ансамблю идентичных генераторов) анодному току  $I_a(t)$ , и *тепловые флуктуации* тока и напряжения в колебательном контуре генератора. Истинным источником этих тепловых флуктуаций является тепловое движение микрочастиц в проводниках, из которых сделан контур. Но в дальнейшем будет обоснована возможность феноменологического описания тепловых флуктуаций в электрической цепи как результата действия некоторой эквивалентной случайной электродвижущей силы (§ 54). Поэтому на схеме автогенератора (рис. 17) в колебательный контур включен «генератор» этой флуктуационной «тепловой» э. д. с.  $e(t)$ .

В обозначениях, ясных из рис. 17, имеем

$$Li + RI - e = \frac{1}{C} \int (I_a + I_{др} - I) dt,$$

$$I_a = Sv_c \left( 1 - \frac{v_g^2}{3V^2} \right), \quad v_c = Mi.$$

Исключая  $I_a$  и  $v_c$ , получаем для тока  $I$  уравнение

$$\ddot{I} + \frac{R}{L} \dot{I} + \frac{I}{LC} = \frac{SM}{LC} \dot{I} \left( 1 - \frac{M^2 \dot{I}^2}{3V^2} \right) + \frac{I_{др}}{LC} + \frac{\dot{e}}{L}.$$

Поскольку мы считаем систему томсоновской, нелинейные и неконсервативные члены уравнения предполагаются малыми — порядка малого параметра  $\mu$ . Обозначим

$$1/LC = \omega_0^2$$

и введем следующие безразмерные переменные и коэффициенты:

$$t' = \omega_0 t, \quad x = I/I_0, \quad I_0 = 2V/\omega_0 M, \quad \omega_0 MS = \mu, \quad \omega_0 (MS - RC) = \mu r.$$

Уравнение приводится тогда к виду

$$\ddot{x} + x = \mu \dot{x} \left( p - \frac{4}{3} \dot{x}^2 \right) + \mu F(t'), \quad (28.1)$$

где

$$\mu F(t') = \frac{M\omega_0}{2V} \left[ I_{др}(t) + \frac{\dot{e}(t)}{L\omega_0} \right] = \frac{1}{I_0} \left[ I_{др}(t) + \frac{\dot{e}(t)}{L\omega_0} \right] \quad (28.2)$$

— действующая на систему флуктуационная «сила», связанная с дробовым током и тепловыми флуктуациями. Эту силу мы тоже считаем малой — того же порядка  $\mu$ , что и диссипативный член уравнения (28.1). Как уже было сказано, ее конкретный вид и вообще ее явное присутствие в уравнении (28.2) понадобятся нам позднее. Сейчас она послужила лишь иллюстрацией того, как могут входить в описание динамической системы случайные воздействия. Дальнейшее развитие теории в этом параграфе, хотя оно и предполагает в действительности определенные ограничения статистических свойств  $F(t')$ , не оперирует с этой случайной «силой» явным образом, т. е. опирается на чисто динамическое уравнение

$$\ddot{x} + x = \mu \dot{x} \left( p - \frac{4}{3} \dot{x}^2 \right). \quad (28.3)$$

К томсоновским системам, близким к гармоническому осциллятору ( $\ddot{x} + x = 0$  при  $\mu = 0$ ), применим, как известно, приближенный (асимптотический) метод Ван-дер-Поля, или, как его называют иначе, метод медленных возмущений. В первом (относительно малого параметра  $\mu$ ) приближении решение уравнения (28.3) отличается от решения для гармонического осциллятора  $x = r \cos(t' + \varphi)$ , где амплитуда  $r$  и фаза  $\varphi$  — постоянные, тем, что  $r$  и  $\varphi$  оказываются *медленными* функциями времени, а именно зависящими от  $t'$  через так называемое «медленное время»  $\theta = \mu t'$ :

$$x = r(\theta) \cos[t' + \varphi(\theta)], \quad (28.4)$$



Таким образом, скорости изменения амплитуды  $r(\theta)$  и фазы  $\varphi(\theta)$  (т. е. добавка  $\dot{\varphi}(\theta)$  к частоте колебаний) являются величинами первого порядка малости относительно параметра  $\mu$ :

$$\dot{r} = \mu r', \quad \dot{\varphi} = \mu \varphi',$$

где штрихом обозначено дифференцирование по медленному времени  $\theta$ .

Напомним упрощенную процедуру получения ван-дер-полевских (или так называемых *укороченных*) уравнений для томсоновской системы, т. е. уравнений первого приближения для медленно меняющихся амплитуды  $r(\theta)$  и фазы  $\varphi(\theta)$ . Излагая эту чисто динамическую теорию, мы будем исходить не из уравнения (28.3) для выбранной модели автогенератора, а из более общего уравнения томсоновской системы с одной степенью свободы:

$$\ddot{x} + x = \mu f(x, \dot{x}) + \mu F(t'). \quad (28.5)$$

Вычисляя с точностью до первого порядка по  $\mu$  производные по  $t'$  искомого решения (28.4), получаем

$$\dot{x} \equiv \frac{dx}{dt'} = -r \sin(t' + \varphi) + \mu [r' \cos(t' + \varphi) - r\varphi' \sin(t' + \varphi)],$$

$$\ddot{x} + x = -2\mu [r' \sin(t' + \varphi) + r\varphi' \cos(t' + \varphi)].$$

Что касается функции  $f(x, \dot{x})$  в (28.5), перед которой уже входит множитель  $\mu$ , то здесь можно ограничиться нулевым приближением для аргумента  $\dot{x}$ . В результате, после сокращения на  $\mu$ , получаем

$$\begin{aligned} -2[r' \sin(t' + \varphi) + r\varphi' \cos(t' + \varphi)] &= \\ &= f[r \cos(t' + \varphi); -r \sin(t' + \varphi)] + F(t'). \end{aligned}$$

Очевидно, функция  $f$ , зависящая от периодических по  $u = t' + \varphi$  аргументов, сама периодична по  $u$  с периодом  $2\pi$  и может быть поэтому разложена в ряд Фурье:

$$f(r \cos u; -r \sin u) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nu + b_n \sin nu),$$

где

$$\left. \begin{matrix} a_n(r) \\ b_n(r) \end{matrix} \right\} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(r \cos u; -r \sin u) \begin{matrix} \cos nu \\ \sin nu \end{matrix} du.$$

Высокая селективность системы (28.5) (неконсервативные члены входят только в порядке  $\mu$ , а в нулевом приближении мы имеем бесконечно селективный осциллятор  $\ddot{x} + x = 0$ ) означает, что из всего ряда Фурье существенны только члены, осциллирую-

щие по  $t'$  с основной безразмерной частотой  $n = 1$ . Постоянную составляющую ( $a_0$ ) и высшие гармоники ( $n = 2, 3, \dots$ ) можно отбросить, так как они далеки от резонанса. Таким образом, уравнение принимает вид

$$\begin{aligned} -2[r' \sin(t' + \varphi) + r\varphi' \cos(t' + \varphi)] = \\ = a_1(r) \cos(t' + \varphi) + b_1(r) \sin(t' + \varphi) + F(t') \end{aligned}$$

По той же причине из всего спектра силы  $F(t')$  играет роль лишь ближайшая окрестность резонансной частоты 1. Вырезая из спектра  $F(t')$  некоторую полосу, центрированную около частоты 1 и имеющую ширину порядка  $\mu$  (но заметно превосходящую ширину резонансной кривой регенерированного контура, рис. 18), мы можем заменить  $F(t')$  процессом  $\tilde{F}(t')$ , у которого

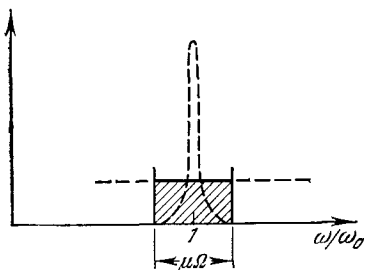


Рис. 18.

спектр ограничен указанной узкой полосой. Такой узкополосный процесс представляет собой колебание с медленно меняющимися амплитудой и частотой (модулированное колебание), и всегда может быть записан в виде

$$\begin{aligned} \tilde{F}(t') = F_{\parallel}(\theta) \cos[t' + \varphi(\theta)] + \\ + \tilde{F}_{\perp}(\theta) \sin[t' + \varphi(\theta)]. \quad (28.6) \end{aligned}$$

В дальнейшем нам придется более обстоятельно заняться свойствами функций  $F_{\parallel}(\theta)$  и  $F_{\perp}(\theta)$ . Здесь же достаточно отметить только то очевидное обстоятельство, что в отсутствие внешней силы [ $F(t') = 0$ ] эти функции, конечно, тоже обращаются в нуль.

Подставляя выражение (28.6) для  $\tilde{F}(t')$  вместо  $F(t')$  в предыдущее уравнение и требуя, чтобы оно тождественно удовлетворялось по  $t'$ , т. е. собирая члены с  $\sin(t' + \varphi)$  и  $\cos(t' + \varphi)$  и приравнивая нулю получающиеся коэффициенты, получаем два уравнения Ван-дер-Поля для  $r$  и  $\varphi$  как функций «медленного времени»  $\theta = \mu t'$ :

$$r' = R(r) - \frac{1}{2} F_{\perp}(\theta), \quad r\varphi' = \Phi(r) - \frac{1}{2} F_{\parallel}(\theta). \quad (28.7)$$

Здесь введены обозначения

$$\left. \begin{aligned} R(r) \\ \Phi(r) \end{aligned} \right\} = -\frac{1}{2} \left\{ \begin{aligned} b_1(r) \\ a_1(r) \end{aligned} \right. = -\frac{1}{2\pi} \int_{-\pi}^{\pi} f(r \cos u; -r \sin u) \begin{aligned} \sin u \\ \cos u \end{aligned} \left. \right\} du. \quad (28.8)$$

Если система автономна [ $F(t') = 0$ ], то уравнения (28.7) будут

$$r' = R(r), \quad r\varphi' = \Phi(r). \quad (28.9)$$

Интегрирование первого из них дает при заданном начальном значении  $r(0)$  закон изменения амплитуды  $r(\theta)$ . Внося это  $r(\theta)$  во второе уравнение, находим при заданном  $\varphi(0)$  закон изменения фазы  $\varphi(\theta)$ . В установившихся режимах (если они существуют) амплитуда  $r$  постоянна ( $r' = 0$ ), а сами эти постоянные значения  $r$  определяются из уравнения  $R(r) = 0$ . Пусть  $r_0$  — какое-либо из решений этого уравнения. Тогда второе уравнение (28.9) определяет соответствующую фазу:

$$\varphi(\theta) = \varphi(0) + \frac{\Phi(r_0)}{r_0} \theta = \varphi(0) + \mu \frac{\Phi(r_0)}{r_0} t'.$$

Таким образом, если  $\Phi(r_0)/r_0 \neq 0$ , автоколебания (28.4) происходят с частотой  $1 + \mu \frac{\Phi(r_0)}{r_0}$ , немного отличной от 1 — собственной частоты линейного осциллятора.

Каковы же уравнения (28.9) для нашей простой модели автотенератора, у которой  $f(x, \dot{x}) = \dot{x} \left( p - \frac{4}{3} \dot{x}^2 \right)$ ? Для функций  $f(x, \dot{x})$  столь простого вида (полином по степеням  $x$  и  $\dot{x}$ , а в данном случае — только по  $\dot{x}$ ) можно даже не вычислять интегралы (28.8), а обойтись элементарной тригонометрией. Подстановка  $\dot{x} = -r \sin u$  дает

$$\dot{x} \left( p - \frac{4}{3} \dot{x}^2 \right) = -pr \sin u + \frac{4}{3} r^3 \sin^3 u = -r(p - r^2) \sin u - \frac{r^3}{3} \sin 3u.$$

Отбросив третью гармонику, имеем, следовательно,  $a_1 = 0$ ,  $b_1 = -r(p - r^2)$  или  $\Phi = 0$ ,  $R = \frac{r}{2}(p - r^2)$ , т. е. получаем уравнения

$$r' = \frac{r}{2}(p - r^2), \quad \varphi' = 0. \quad (28.10)$$

Возможные состояния равновесия — это  $r = 0$  и при  $p > 0$  еще  $r = \sqrt{p}$ . Элементарное исследование показывает, что при  $p < 0$  состояние равновесия  $r = 0$  устойчиво, а при  $p > 0$  неустойчиво, но устойчив предельный цикл с  $r = \sqrt{p}$ . Фаза  $\varphi = \varphi_0 = \text{const}$ , т. е. сохраняет начальное значение. На рис. 19 показан ход траекторий на фазовой плоскости  $x = r \cos(t' + \varphi)$ ,  $\dot{x} \approx -r \sin(t' + \varphi)$ , а в нижней части рисунка — соответствующая картина на плоскости Ван-дер-Поля, т. е. на плоскости с полярными координатами  $r$  и  $\varphi$ . Связь обеих картин в рассматриваемом первом приближении очень проста: если вращать плоскость Ван-дер-Поля вокруг начала отсчета с угловой частотой

единица, то медленные движения изображающей точки по лучу  $\varphi = \varphi_0 = \text{const}$  перейдут в спирали, а устойчивое положение равновесия  $r = \sqrt{\rho}$  — в окружность (предельный цикл) на фазовой плоскости  $(x, \dot{x})$ . Таково в общих чертах решение динамической задачи.

Как изменятся эти результаты при учете случайных воздействий и чем эти воздействия обусловлены? Остановимся сначала на втором вопросе.

Анодный ток лампы испытывает флуктуации не только из-за дробового эффекта (§§ 5, 10), но и вследствие «эффекта мерцания» — хаотических вариаций эмиссии катода. В самом коле-

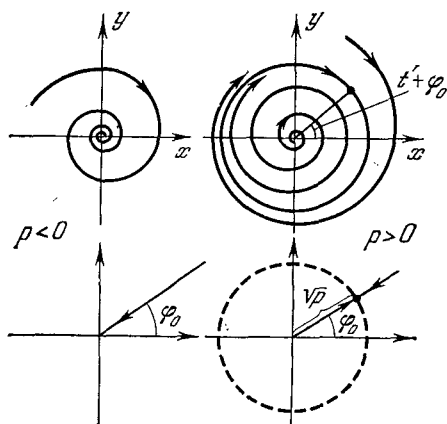


Рис. 19.

бательном контуре флуктуации тока и напряжения обусловлены не только тепловым движением электронов, но и рядом других причин. Генератор подвержен разнообразным внешним воздействиям — механическим, которые вызывают, в частности, вибрацию электродов лампы (микрофонный эффект); температурным, которые могут влиять на значения параметров схемы; электромагнитным (электрические флуктуации в источниках питания, наводки) и т. п. Часть этих воздействий в принципе

устраняема: можно хорошо заблокировать источники питания, можно тщательно заэкранировать самый генератор, термостатировать и амортизировать его и т. д. Но такие явления, как дробовой эффект и тепловые флуктуации, принципиально неустранимы, так как они связаны с атомистической структурой электрического заряда. Эффект мерцания, обусловленный довольно сложными процессами диффузии и адсорбции на поверхностном слое катода, присутствует, как показывает опыт, во всех случаях, но, в отличие от дробовых и тепловых шумов, протекает довольно медленно по сравнению с высокочастотными колебаниями генератора.

Можно подразделить случайные воздействия на два вида. Одни вызывают медленные, но значительные изменения параметров устройства и соответствующие уходы частоты. Г. С. Горелик [16] предложил называть эти уходы и связанную с ними немонохроматичность колебаний *техническими*, подчеркивая этим их обусловленность внешними факторами, от которых

в принципе можно избавиться (хотя эффект мерцания дает вклад именно в эту группу). Принципиально же неустранимую немонохроматичность, обусловленную чрезвычайно быстрой, но неглубокой хаотической модуляцией вследствие дробовых и тепловых флуктуаций, он предложил называть *естественной*. В дальнейшем мы еще вернемся к вопросу о влиянии обоих названных видов хаотической модуляции на спектр автоколебаний. В данный момент нас интересует протекание этих явлений во времени и возможность описания их как *непрерывного* марковского процесса.

Как уже отмечалось (§ 26), такое описание предполагает достаточную «густоту» независимых случайных толчков: вероятность конечных изменений состояния за макроскопически малое время  $\Delta t'$  должна стремиться к нулю быстрее, чем  $\Delta t'$ , но вместе с тем интервал  $\Delta t'$  должен быть велик по сравнению со временем между случайными толчками, т. е. за время  $\Delta t'$  система должна испытывать очень много независимых случайных воздействий.

Что является макроскопическим масштабом малости  $\Delta t'$ ? Очевидно, *период* автоколебаний (равный  $2\pi$  по безразмерному времени  $t'$ ), так как другие временные характеристики томсоновской системы (время установления генератора или отдельно взятого колебательного контура) охватывают очень много ( $\sim 1/\mu$ ) периодов. Следовательно,  $2\pi \gg \Delta t'$ , которое в свою очередь гораздо больше времени между толчками. Но этому условию удовлетворяют только дробовые и тепловые флуктуации. Случайные воздействия, с которыми связаны технические уходы частоты, гораздо более медленны по сравнению не только с периодом, но даже со временем установления генератора. Отсюда ясно, что они не могут быть включены в схему случайного процесса без вероятностного последствия, если, конечно, не отказываться от трактовки генератора как системы с *одной* степенью свободы. Таким образом, понимая под состоянием совокупность двух переменных  $x$  и  $\dot{x}$  (или  $r$  и  $\varphi$ ) в момент времени  $t'$  и описывая случайные изменения состояния как непрерывный двумерный марковский процесс, мы можем учесть влияние только *дробовых* и *тепловых* флуктуаций.

В результате этих флуктуаций, быстро меняющихся на протяжении даже маленького отрезка предельного цикла (малой доли периода), изображающая точка на фазовой плоскости ( $x, y$ ) [или на плоскости ( $r, \varphi$ ), где систематическое обращение по часовой стрелке с периодом  $2\pi$  исключено] совершает своего рода брауновское движение. Для того чтобы написать соответствующее уравнение Эйнштейна — Фоккера для вероятности перехода  $v(\theta, r, \varphi | \theta_0, r_0, \varphi_0)$ , необходимо конкретизировать вид коэффициентов  $B_{ik}$ . Мы допустим, что случайные толчки

изотропны, т. е. справедливы выражения (27.9):

$$B_{rr} = r^2 B_{\varphi\varphi} = B, \quad B_{r\varphi} = 0.$$

Кроме того, мы примем, что толчки *однородны*, т. е.  $B$  — постоянная величина, не зависящая от состояния генератора  $(r, \varphi)$ .

Заметим, что это последнее предположение, строго говоря, непригодно для лампового генератора. Оно справедливо, если речь идет о тепловых флуктуациях в колебательном контуре, так как их интенсивность не зависит от силы текущего по проводникам тока. Но интенсивность флуктуаций анодного тока лампы (дробовой эффект) прямо пропорциональна среднему значению тока [см. (10.9)]. При наличии автоколебаний средний ток периодически пульсирует в такт с колебаниями напряжения на сетке. Следовательно, в переменных  $(r, \varphi)$  коэффициент  $B$  должен был бы зависеть не только от  $r$  и  $\varphi$ , но и от времени  $t'$  (периодическая нестационарность при установившемся режиме). Допуская постоянство  $B$ , мы переходим к сильно упрощенной модели, но она все же позволяет составить известное представление о влиянии флуктуаций на поведение генератора и довольно близка к реальным условиям в тех случаях, когда колебания анодного тока лампы относительно невелики.

Конечно, изложенные качественные соображения нельзя считать *доказательством* того, что при воздействии только «густых» дробовых и тепловых флуктуаций случайный процесс  $(r, \varphi)$  в генераторе действительно является непрерывным марковским процессом и подчинен (при дополнительном допущении однородности и изотропности случайных толчков) уравнению Эйнштейна — Фоккера (27.12). Более того, в это уравнение входят, кроме коэффициента  $B$ , еще детерминированные функции  $R(r, \varphi)$  и  $\Phi(r, \varphi)$ , которые надо откуда-то взять. У нас пока нет регулярного способа нахождения этих функций, так что и в этом вопросе мы тоже пока вынуждены опереться на наводящие соображения.

В § 26 мы убедились, что в случае одномерного марковского процесса коэффициент  $A(x, t)$  в уравнении Эйнштейна — Фоккера совпадает с правой частью «динамического» уравнения (26.13) с точностью до аддитивной случайной функции, среднее значение которой равно нулю. В нашей двумерной задаче такие функции могли бы появиться в правых частях динамических уравнений (28.10), если бы при их выводе мы не отбросили в исходном уравнении (28.1) случайную «силу»  $F(t)$  (у которой среднее значение как раз равно нулю). Естественно допустить, что в качестве  $R$  и  $\Phi$  в (27.12) надо подставить правые части динамических уравнений (28.10), т. е. положить

$$R = \frac{r}{2} (p - r^2), \quad \Phi = 0. \quad (28.11)$$

Тогда уравнение (27.12) для вероятности перехода  $v(r, \varphi | \theta - \theta_0, r_0, \varphi_0)$  будет

$$\frac{\partial v}{\partial t} = -\frac{\partial Rv}{\partial r} + \frac{B}{2} \left[ \frac{\partial}{\partial r} \left( r \frac{\partial}{\partial r} \left( \frac{v}{r} \right) \right) + \frac{1}{r^2} \frac{\partial^2 v}{\partial \varphi^2} \right]. \quad (28.12)$$

Однородность процесса во времени ( $v$  зависит от  $\theta - \theta_0$ ) является следствием независимости  $R$  и  $B$  от  $\theta$ .

Все наши правдоподобные допущения, приведшие к уравнению (28.12), будут обоснованы в следующей главе, посвященной стохастическим дифференциальным уравнениям.

Физически очевидно, что процесс эргодичен — система обладает конечной «памятью», так что с ростом  $\theta - \theta_0$  вероятность перехода превращается в одномерную вероятность состояния  $w_1$ , не зависящую от начальных значений  $r_0$  и  $\varphi_0$ :

$$v(r, \varphi | \theta - \theta_0, r_0, \varphi_0) \rightarrow w_1(r, \varphi), \quad |\theta - \theta_0| \rightarrow \infty. \quad (28.13)$$

Разумеется, полное решение уравнения (28.12) при начальном условии

$$v(r, \varphi | 0, r_0, \varphi_0) = \delta(r - r_0) \delta(\varphi - \varphi_0) \quad (28.14)$$

обнаружило бы справедливость (28.13) без каких-либо добавочных предположений, но мы сразу ограничимся более простой задачей — получением *стационарного* решения  $w_1(r, \varphi)$ . Оно удовлетворяет уравнению

$$\frac{\partial R w_1}{\partial r} - \frac{B}{2} \left[ \frac{\partial}{\partial r} \left( r \frac{\partial}{\partial r} \left( \frac{w_1}{r} \right) \right) + \frac{1}{r^2} \frac{\partial^2 w_1}{\partial \varphi^2} \right] = 0.$$

Подчеркнем теперь одно обстоятельство, которое станет совершенно очевидным в последующем (§ 29): стационарное распределение возможно только при условии, что в качестве области изменения фазы  $\varphi$  берется *конечный* интервал  $(0, 2\pi)$ . Это вовсе не единственно возможная постановка задачи, но если она принята, то без дальнейшего ясно, что все значения  $\varphi$  в интервале  $(0, 2\pi)$  равноправны. Это значит, что распределение по  $\varphi$  равномерно ( $w_1$  не зависит от  $\varphi$ ), так что уравнение для  $w_1$  принимает вид

$$\frac{d}{dr} \left( R w_1 - \frac{B}{2} r \frac{d}{dr} \left( \frac{w_1}{r} \right) \right) = 0.$$

Нетрудно убедиться, что не обладает особенностью в точке  $r = 0$  только то решение, которое обращает поток  $w_1$ , т. е. выражение, стоящее в скобках, в нуль. Интегрируя вторично, находим

$$w_1 = C r \exp \left\{ \frac{2}{B} \int_0^r R dr \right\},$$

причем  $C$  определяется из условия нормировки

$$\int_0^{2\pi} d\varphi \int_0^{\infty} w_1 dr = 2\pi C \int_0^{\infty} \exp \left\{ \frac{2}{B} \int_0^r R dr \right\} r dr = 1.$$

Подставив в эти выражения  $R$  из (28.10) и вычислив интегралы, получаем искомое стационарное распределение:

$$w_1(r, \varphi) dr d\varphi = \frac{e^{-(p-r^2)^2/4B}}{\pi \sqrt{\pi B} [1 + \operatorname{erf}(p/2\sqrt{B})]} r dr d\varphi, \quad (28.15)$$

где

$$\operatorname{erf} x = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$$

На рис. 20 изображена поверхность  $w_1(r, \varphi)$  при  $p > 0$  (возбужденный генератор), когда эта поверхность имеет вид кратера. Наивероятнейшее значение  $r$  есть

$$r_m = \sqrt{\frac{p}{2} + \sqrt{\frac{p^2}{4} + B}},$$

т. е.

$$r_m \approx \sqrt{p} \left( 1 + \frac{B}{2p^2} \right) \quad \text{при } p \gg 2\sqrt{B},$$

$$r_m \approx \sqrt{\frac{B}{|p|}} \quad \text{при } p \ll -2\sqrt{B}.$$

В последнем случае, соответствующем достаточному удалению от порога самовозбуждения в сторону отрицательных  $p$ , можно пренебречь в показателе (28.15) четвертой степенью  $r$  и воспользоваться асимптотическим разложением

$$\operatorname{erf}(-x) = -1 + \frac{e^{-x^2}}{\sqrt{\pi}|x|} - \dots$$

Мы получаем тогда

$$w_1(r, \varphi) dr d\varphi = \frac{|p|}{B} e^{-r^2|p|/2B} r dr \cdot \frac{d\varphi}{2\pi}, \quad (28.16)$$

т. е. для амплитуды  $r$ , как этого и следовало ожидать в случае линейного диссипативного контура, имеет место распределение Релея (24.10), причем  $2\sigma^2 = \overline{r^2} = 2B/|p|$ .

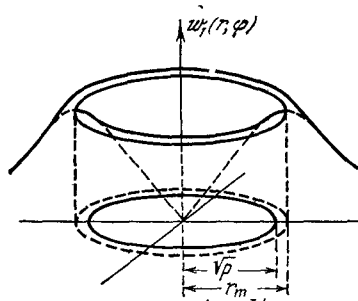


Рис. 20.



Напротив, если генератор достаточно сильно возбужден ( $p \gg 2\sqrt{B}$ ), то «вал», окружающий «кратер», имеет заметную высоту лишь над кольцом, содержащим предельный цикл. Полагая  $r = \sqrt{p} + \rho$ , ограничиваясь квадратом  $\rho$  в показателе (28.15) и пользуясь тем, что теперь

$$\operatorname{erf}(p/2\sqrt{B}) \approx 1,$$

получаем

$$\begin{aligned} \omega_1(\rho, \varphi) d\rho d\varphi = \\ = \sqrt{\frac{p}{\pi B}} e^{-\rho^2 p/B} d\rho \frac{d\varphi}{2\pi}. \end{aligned} \quad (28.17)$$

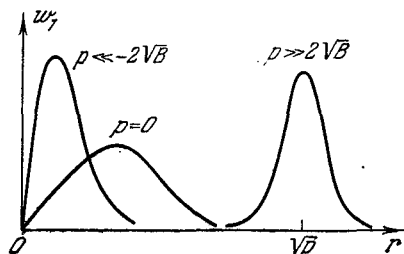


Рис. 21.

Таким образом, радиальное отклонение  $\rho$  от предельного цикла распределено в этом случае нормально с дисперсией  $\sigma^2 = \rho^2 = B/2p$ .

На рис. 21 показан ход  $\omega_1$  в функции от  $r$  для трех значений  $p$ , в том числе на границе самовозбуждения  $p = 0$ . Разумеется, при всех значениях  $p$  увеличение интенсивности случайных толчков (рост  $B$ ) ведет к расплыванию максимума  $\omega_1$ .

## § 29. Флуктуации при больших амплитудах автоколебаний

Рассмотрим подробнее случай больших положительных значений  $p$ . Условие  $p \gg 2\sqrt{B}$  означает, очевидно, что радиус предельного цикла  $\sqrt{p}$  велик по сравнению со стандартом флуктуаций амплитуды:

$$\sqrt{p} \gg \sqrt{B/2p} = \sqrt{\rho^2}.$$

Это позволяет линейаризовать динамические уравнения (28.10) относительно  $\rho = r - \sqrt{p}$ . Они принимают вид

$$\rho' = -p\rho, \quad \varphi' = 0, \quad (29.1)$$

т. е. отклонение изображающей точки от предельного цикла ведет себя так, как если бы эта точка была привязана к циклу пружиной и двигалась в среде с вязким трением. Инкремент  $p$  играет роль отношения коэффициента упругости пружины к коэффициенту вязкого трения и характеризует «прочность» предельного цикла — скорость возвращения изображающей точки на цикл после начального отклонения от него.

Соответственно (29.1) упрощается и уравнение Эйнштейна — Фоккера (28.12), в котором надо положить  $\partial/\partial r = \partial/\partial \rho$ ,  $r = \sqrt{p}$

и  $R = -\rho r$ :

$$\frac{\partial v}{\partial \theta} = \frac{\partial \rho}{\partial \rho} \frac{\partial v}{\partial \rho} + \frac{B}{2} \left( \frac{\partial^2 v}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial^2 v}{\partial \varphi^2} \right). \quad (29.2)$$

Упрощение настолько значительно, что теперь можно не ограничиваться отысканием стационарного распределения, а получить решение (29.2) при произвольных начальных условиях  $\rho = \rho_0$ ,  $\varphi = \varphi_0$ ,  $\theta = \theta_0$  (конечно, надо придерживаться при этом условия  $\rho_0 \ll \sqrt{p}$ ). Для вероятности перехода  $v(\rho, \varphi | \theta - \theta_0, \rho_0, \varphi_0)$  указанные начальные условия означают, что

$$v(\rho, \varphi | 0, \rho_0, \varphi_0) = \delta(\rho - \rho_0) \delta(\varphi - \varphi_0). \quad (29.3)$$

Если вновь предположить, что возможные значения  $\varphi$  лежат в интервале  $(0, 2\pi)$ , т. е. искать решение (29.2), периодическое по  $\varphi$  с периодом  $2\pi$ , то обычная процедура разделения переменных приводит к следующему результату:

$$v(\rho, \varphi | \theta - \theta_0, \rho_0, \varphi_0) = v_\rho(\rho | \theta - \theta_0, \rho_0) v_\varphi(\varphi | \theta - \theta_0, \varphi_0), \quad (29.4)$$

т. е. в любой момент времени  $\rho$  и  $\varphi$  независимы, причем флуктуации амплитуды  $\rho$  распределены нормально:

$$v_\rho(\rho | \theta - \theta_0, \rho_0) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(\rho - \bar{\rho})^2/2\sigma^2}, \quad (29.5)$$

со средним значением и дисперсией, равными соответственно

$$\bar{\rho} = \rho_0 e^{-p(\theta - \theta_0)}, \quad \sigma^2 = \frac{B}{2p} [1 - e^{-2p(\theta - \theta_0)}], \quad (29.6)$$

а распределение фазы  $\varphi$  есть

$$v_\varphi(\varphi | \theta - \theta_0, \varphi_0) = \frac{1}{2\pi} \left\{ 1 + 2 \sum_{n=1}^{\infty} e^{-n^2 D(\theta - \theta_0)} \cos n(\varphi - \varphi_0) \right\}, \quad (29.7)$$

где  $D = B/2p$ .

В начальный момент  $\theta = \theta_0$  имеем  $\bar{\rho} = \rho_0$ ,  $\sigma^2 = 0$ , т. е. распределение  $\rho$  превращается в  $\delta(\rho - \rho_0)$ , а распределение  $\varphi$  при  $\theta = \theta_0$  дает  $\delta(\varphi - \varphi_0)$ , поскольку  $1 + 2 \sum_{n=1}^{\infty} \cos n(\varphi - \varphi_0)$  есть разложение Фурье для периодически-повторяющейся функции  $2\pi\delta(\varphi - \varphi_0)$ . Как для амплитуды, так и для фазы имеет место эргодичность: при  $\theta \rightarrow \infty$  устанавливаются не зависящие от начальных условий стационарные распределения, а именно — нормальное распределение для  $\rho$  с  $\bar{\rho} = 0$  и дисперсией  $\sigma^2 = B/2p$  и равномерное в интервале  $(0, 2\pi)$  распределение для  $\varphi$ , что мы и получили ранее [см. (28.17)].

На рис. 22 показано, как устанавливается распределение амплитуды: среднее значение  $\bar{\rho}$  затухает по такому же экспоненциальному закону, по какому изображающая точка приближалась бы к предельному циклу в чисто динамической задаче. Наличие случайных толчков ведет вместе с тем к нарастанию дисперсии  $(\rho - \bar{\rho})^2$  вплоть до ее установившегося значения. Что

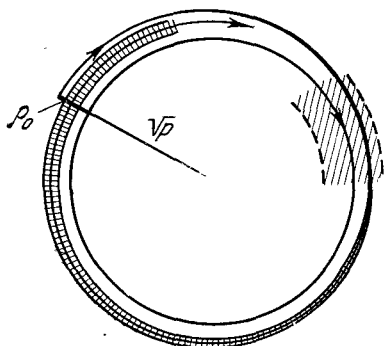


Рис. 22.

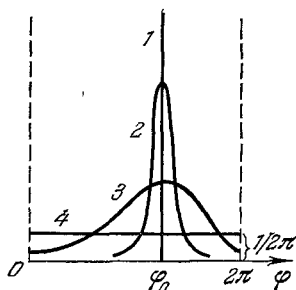


Рис. 23.

касается распределения фазы, то оно меняется так, как изображено на рис. 23: ансамбль идентичных генераторов, запущенных в момент  $\theta = \theta_0$  с одинаковой фазой  $\varphi_0$ , диффузионно растекается по предельному циклу, и спустя достаточно долгое время экземпляры из этого ансамбля населяют весь цикл равномерно.

Заметим теперь следующее. Выбор интервала  $(0, 2\pi)$  для  $\varphi$  означает применительно к отдельному генератору, что из полного набега его фазы мы сбрасываем всякое целое число полных циклов. Но если речь идет об использовании автогенератора в качестве часов, то отнюдь не безразлично, произошел ли уход фазы на угол  $\alpha$ , или на  $\alpha + 2\pi$ ,  $\alpha + 4\pi$  и т. д. Другими словами, нас здесь интересует именно *полный* набег фазы, так что для ее возможных значений надо брать интервал от  $-\infty$  до  $+\infty$ . В этом случае решение уравнения (29.2) будет по-прежнему иметь вид (29.4); для флуктуаций амплитуды останутся в силе формулы (29.5) и (29.6), но вероятность перехода для фазы  $v_\varphi(\varphi | \theta - \theta_0, \varphi_0)$  не будет периодична по  $\varphi$  и при начальном условии  $\delta(\varphi - \varphi_0)$  выразится, как нетрудно убедиться, формулой ( $D = B/2p$ )

$$v_\varphi(\varphi | \theta - \theta_0, \varphi_0) = \frac{1}{\sqrt{4\pi D(\theta - \theta_0)}} \exp \left\{ -\frac{(\varphi - \varphi_0)^2}{4D(\theta - \theta_0)} \right\}. \quad (29.8)$$

Мы возвращаемся, таким образом, к случаю изотропных блужданий брауновской частицы по бесконечной прямой (см. §§ 4,

7, 24) и, соответственно, уже не имеем эргодичности: не зависящее от  $\varphi_0$  стационарное распределение не достигается ни при каком  $\theta$ , и средний квадрат флуктуационного набег фазы за промежутков времени  $\theta - \theta_0$  неограниченно растет пропорционально  $\theta - \theta_0$ :

$$\overline{(\varphi - \varphi_0)^2} = 2D(\theta - \theta_0). \quad (29.9)$$

Связь обеих постановок задачи очень проста (рис. 24): если неограниченно растекающееся по бесконечной оси  $\varphi$  распределение (29.8) «свернуть» путем

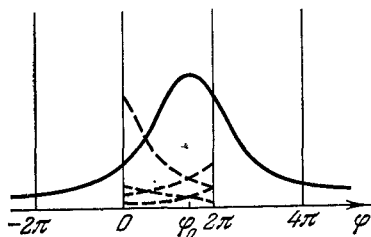


Рис. 24.

переноса всех интервалов  $[2\pi n, 2\pi(n+1)]$  в полосу  $(0, 2\pi)$ , то в этой полосе получится распределение (29.7), т. е. мы вернемся к картине установления, показанной на рис. 23. Очевидно, до тех пор, пока расплывание распределения (29.7) еще невелико, т. е.  $\overline{(\varphi - \varphi_0)^2} \ll 2\pi$ , законы распределения (29.7) и (29.8) прак-

тически совпадают в интервале  $(0, 2\pi)$ . В этом нетрудно убедиться и прямым расчетом, заменив сумму в (29.7) интегралом.

Как следствие предположения об однородности случайных толчков рассматриваемый марковский процесс однороден по времени: вероятности перехода (29.5), (29.7) и (29.8) зависят лишь от разности между конечным ( $\theta$ ) и начальным ( $\theta_0$ ) моментами. Возьмем в качестве начального момента времени  $\theta = \mu t'$ , а в качестве конечного — момент  $\theta + \tau = \mu(t' + \tau')$ . Вероятности перехода для  $\rho$  и  $\varphi$  будут зависеть только от  $\tau = \mu\tau'$ . Обозначив значения  $\rho$  и  $\varphi$  в момент  $\theta + \tau$  через  $\rho_\tau$  и  $\varphi_\tau$ , можно переписать (29.5) и (29.7) в виде

$$v_\rho(\rho_\tau | \tau, \rho) = \frac{1}{\sqrt{2\pi} \sigma_\tau} e^{-(\rho_\tau - \bar{\rho}_\tau)^2 / 2\sigma_\tau^2},$$

$$\bar{\rho}_\tau = \rho e^{-p\tau}, \quad \sigma_\tau^2 = \frac{B}{2p} (1 - e^{-2p\tau}), \quad (29.10)$$

$$v_\varphi(\varphi_\tau | \tau, \varphi) = \frac{1}{2\pi} \left\{ 1 + 2 \sum_{n=1}^{\infty} e^{-n^2 D \tau} \cos n(\varphi_\tau - \varphi) \right\}.$$

При  $\tau \rightarrow \infty$  отсюда получаются стационарные одномерные распределения:

$$\omega_{1\rho}(\rho) = \frac{1}{\sqrt{2\pi} \sigma_\infty} e^{-\rho^2 / 2\sigma_\infty^2}, \quad \sigma_\infty^2 = \frac{B}{2p} = 0, \quad (29.11)$$

$$\omega_{1\varphi}(\varphi) = \frac{1}{2\pi}. \quad (29.12)$$

Как мы помним (§ 15), вероятность перехода и одномерная вероятность — это все, что необходимо для того, чтобы написать любое  $n$ -мерное распределение марковского процесса. В частности, двумерные распределения  $\rho$  и  $\varphi$  будут

$$\begin{aligned} w_{2\rho}(\rho_\tau, \tau, \rho) &= w_{1\rho}(\rho) v_\rho(\rho_\tau | \tau, \rho) = \\ &= \frac{1}{2\pi\sigma_\infty\sigma_\tau} \exp \left\{ -\frac{\rho^2}{2\sigma_\infty^2} - \frac{(\rho_\tau - \bar{\rho}_\tau)^2}{2\sigma_\tau^2} \right\}, \quad (29.13) \end{aligned}$$

$$\begin{aligned} w_{2\varphi}(\varphi_\tau, \tau, \varphi) &= w_{1\varphi}(\varphi) v_\varphi(\varphi_\tau | \tau, \varphi) = \\ &= \frac{1}{4\pi^2} \left\{ 1 + 2 \sum_{n=1}^{\infty} e^{-n^2 D\tau} \cos n(\varphi_\tau - \varphi) \right\}. \quad (29.14) \end{aligned}$$

Поскольку  $w_{1\rho}(\rho)$  и  $w_{1\varphi}(\varphi)$  не зависят от  $\theta$ , а вероятности перехода однородны по  $\theta$ , рассматриваемый марковский процесс *стационарен*.

Располагая двумерными распределениями (29.13) и (29.14), можно вычислить средние значения каких-либо функций от  $\rho$ ,  $\rho_\tau$  и от  $\varphi$ ,  $\varphi_\tau$ . Мы воспользуемся этим для расчета функции корреляции исследуемого автоколебательного процесса

$$x(t') = r(\theta) \cos[t' + \varphi(\theta)] = (\sqrt{p} + \rho) \cos(t' + \varphi),$$

которая понадобится в дальнейшем при изучении спектра автогенератора. Интересующая нас функция корреляции есть

$$\psi_x = \overline{xx_{\tau'}} - \bar{x}\bar{x}_{\tau'},$$

где  $x \equiv x(t')$  и  $x_{\tau'} \equiv x(t' + \tau') = (\sqrt{p} + \rho_\tau) \cos(t' + \tau' + \varphi_\tau)$ .

Поскольку  $\rho$  и  $\varphi$  в любой момент времени независимы, а одномерное распределение фазы (29.12) равномерно в интервале  $(0, 2\pi)$ , имеем  $\cos \varphi = \sin \varphi = 0$ . Следовательно, для всякого  $t'$

$$\bar{x} = (\sqrt{p} + \bar{\rho}) \overline{\cos(t' + \varphi)} = 0.$$

Учитывая, далее, что, согласно (29.11),  $\bar{\rho} = \bar{\rho}_\tau = 0$ , получаем

$$\begin{aligned} \psi_x &= \overline{xx_{\tau'}} = \overline{(\sqrt{p} + \rho)(\sqrt{p} + \rho_\tau) \cos(t' + \varphi) \cos(t' + \tau' + \varphi_\tau)} = \\ &= \frac{1}{2} (p + \overline{\rho\rho_\tau}) \{ \overline{\cos(\tau' + \varphi_\tau - \varphi)} + \overline{\cos(2t' + \tau' + \varphi_\tau + \varphi)} \}. \quad (29.15) \end{aligned}$$

Таким образом, для нахождения  $\psi_x$  надо вычислить следующие средние:

$$\overline{\rho\rho_\tau}, \quad \overline{\frac{\sin}{\cos}} \{ (\varphi_\tau - \varphi) \}, \quad \overline{\frac{\sin}{\cos}} \{ (\varphi_\tau + \varphi) \}.$$

В соответствии с (29.13) имеем

$$\begin{aligned}\overline{\rho\rho_\tau} &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \rho\rho_\tau w_{2\rho}(\rho_\tau, \tau, \rho) d\rho d\rho_\tau = \\ &= \int_{-\infty}^{+\infty} \rho w_{1\rho}(\rho) d\rho \int_{-\infty}^{+\infty} \rho_\tau v_\rho(\rho_\tau | \tau, \rho) d\rho_\tau.\end{aligned}$$

Внутренний интеграл — это условное среднее значение  $\bar{\rho}_\tau$ , равное  $\rho e^{-\rho\tau}$  [см. (29.10)]. Следовательно [см. (29.11)],

$$\overline{\rho\rho_\tau} = e^{-\rho\tau} \int_{-\infty}^{+\infty} \rho^2 w_{1\rho}(\rho) d\rho = \overline{\rho^2} e^{-\rho\tau} = \sigma_\infty^2 e^{-\rho\tau} = \frac{B}{2\rho} e^{-\rho\tau}.$$

Учитывая, что  $\overline{\rho\rho_{-\tau}} = \overline{\rho\rho_\tau}$ , можно записать этот результат при произвольном знаке  $\tau$  в виде

$$\overline{\rho\rho_\tau} = \frac{B}{2\rho} e^{-\rho|\tau|}. \quad (29.16)$$

Из того, что  $w_{2\varphi}(\varphi_\tau, \tau, \varphi)$  — четная функция  $\varphi_\tau - \varphi$  [см. (29.14)], тотчас же следует, что

$$\overline{\cos(\varphi_\tau + \varphi)} = \overline{\sin(\varphi_\tau + \varphi)} = \overline{\sin(\varphi_\tau - \varphi)} = 0. \quad (29.17)$$

Остается вычислить  $\overline{\cos(\varphi_\tau - \varphi)}$ , т. е. интеграл [см. (29.14)]

$$\begin{aligned}\overline{\cos(\varphi_\tau - \varphi)} &= \int_0^{2\pi} \int_0^{2\pi} \cos(\varphi_\tau - \varphi) w_{2\varphi}(\varphi_\tau, \tau, \varphi) d\varphi d\varphi_\tau = \\ &= \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} \cos(\varphi_\tau - \varphi) \left\{ 1 + 2 \sum_{n=1}^{\infty} e^{-n^2 D\tau} \cos n(\varphi_\tau - \varphi) \right\} d\varphi d\varphi_\tau.\end{aligned}$$

Очевидно, отличный от нуля результат дает только член суммы с  $n = 1$ :

$$\overline{\cos(\varphi_\tau - \varphi)} = \frac{e^{-D\tau}}{2\pi^2} \int_0^{2\pi} \int_0^{2\pi} \cos^2(\varphi_\tau - \varphi) d\varphi d\varphi_\tau = e^{-D\tau},$$

или при произвольном знаке  $\tau$ :

$$\overline{\cos(\varphi_\tau - \varphi)} = e^{-D|\tau|}. \quad (29.18)$$

Внося (29.16) — (29.18) в (29.15), получаем

$$\psi_x(\tau') = \frac{1}{2} (\rho + D e^{-\rho|\tau'|}) e^{-D|\tau'|} \cos \tau' \quad (\tau = \mu\tau', D = \frac{B}{2\rho}). \quad (29.19)$$

Три сомножителя, входящие в  $\psi_x(\tau')$ , меняются в функции от  $\tau'$  с существенно различными скоростями. Наиболее быстро — с частотой автоколебаний — осциллирует  $\cos \tau'$ . Гораздо медленнее меняется экспонента

$$e^{-P|\tau|} = e^{-\mu P|\tau'|},$$

вошедшая через функцию корреляции амплитудных флуктуаций (29.16). Последняя уменьшается в  $e$  раз за время  $1/\mu P$ , причем в обычных условиях безразмерный инкремент  $\mu P$  — величина порядка  $10^{-2}$ — $10^{-3}$ . Еще медленнее убывает экспонента

$$e^{-D|\tau|} = e^{-\mu D|\tau'|},$$

где  $2\mu D$  — коэффициент диффузии фазы:

$$\overline{(\varphi_\tau - \varphi)^2} = 2D\tau = 2\mu D\tau'. \quad (29.20)$$

Для диффузии, обусловленной естественными флуктуациями, величина  $2\mu D$  имеет порядок  $10^{-10}$ — $10^{-13}$ . Это значит, что средний квадратичный уход фазы примерно на  $\pi$  занимает по безразмерному времени промежуток

$$\tau' \approx \frac{\pi^2}{2\mu D},$$

а по обычному времени

$$\Delta t = \frac{\tau'}{\omega_0} \approx \frac{\pi}{4\mu D f_0} \approx (10^{10} - 10^{13}) \frac{1}{f_0}.$$

При  $f_0 = 10$  Мгц это составляет  $\Delta t' \approx 10^3$ — $10^6$  сек.

Хорошо известно, однако, что расхождение на  $\pi$  между фазами двух специально не стабилизированных генераторов происходит на таких частотах несравненно быстрее — за немногие секунды. Это связано с наличием *технических* уходов фазы (§ 28). Казалось бы, технические уходы должны полностью маскировать естественную диффузию фазы, но подобное заключение было бы слишком поспешным, так как технические уходы фундаментально отличаются от естественных своей гораздо более длительной корреляцией. Мы вернемся к этому вопросу в дальнейшем (§ 53).

Заметим в заключение этого параграфа, что параметр  $B$ , присутствующий во всех предыдущих формулах и определяющий, в частности, коэффициент диффузии фазы, вошел в эти формулы из исходного уравнения Эйнштейна — Фоккера (28.12). Мы не имеем пока никакой связи  $B$  с величинами, описывающими конкретный механизм случайных воздействий на генератор (дробовой ток, тепловой шум), или, говоря более определенно, со статистическими характеристиками случайной силы (28.2). Это еще один пробел развитой здесь чисто *вероятностной* схемы, который нам тоже предстоит восполнить в следующей главе.

### § 30. Вращательное брауновское движение. Случайная рефракция луча

Пусть сферическая частица взвешена в жидкости или газе, и пусть с частицей жестко связана некоторая ось, направление которой, характеризующее ориентацию частицы, определяется полярными углами  $\theta$  и  $\varphi$ . Отвлечемся от поступательного движения частицы и от ее вращения вокруг оси. Тогда под действием ударов молекул или — что при данной постановке задачи то же самое — под действием флуктуаций тангенциальных (вязких) напряжений частица (ее ось) будет случайным образом поворачиваться, совершая так называемое вращательное брауновское движение<sup>1)</sup>. Это означает, что точка  $(\theta, \varphi)$  совершает поступательное брауновское движение на поверхности единичной сферы.

Очевидно, случайный вращающий момент будет изотропен (все его направления равноправны) и однороден (интенсивность толчков не зависит от ориентации частицы). Уравнение Эйнштейна — Фоккера, отвечающее этим условиям и определяющее скалярную плотность вероятности перехода  $\tilde{v}(\theta, \varphi | t - t_0, \theta_0, \varphi_0)$ , т. е. такую плотность, что  $dP = \tilde{v} \sin \theta d\theta d\varphi$ , уже было написано [см. (27.17)]. Оно упрощается, если полярный угол  $\theta$  отсчитывается от начального направления оси частицы ( $\theta_0 = 0$ ). Вследствие симметрии распределение по азимуту  $\varphi$  становится тогда равномерным, а уравнение для вероятности перехода  $\tilde{v}(\theta | t - t_0, 0)$  принимает вид (27.17):

$$\sin \theta \frac{\partial \tilde{v}}{\partial t} = \frac{B}{2} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial \tilde{v}}{\partial \theta} \right), \quad \tilde{v}(\theta | 0, 0) = \frac{\delta(\theta)}{2\pi \sin \theta}. \quad (30.1)$$

Решение уравнения (30.1) с этим начальным условием есть  $v(\theta | t - t_0, 0) =$

$$= \frac{1}{4\pi} \sum_{n=0}^{\infty} (2n+1) P_n(\cos \theta) \exp \left\{ -\frac{n(n+1)}{2} B(t-t_0) \right\}, \quad (30.2)$$

где  $P_n(\cos \theta)$  — полиномы Лежандра. При  $t \rightarrow \infty$  достигается стационарное распределение: из всей суммы остается только член с  $n = 0$ , так что

$$\lim_{t \rightarrow \infty} \tilde{v}(\theta | t - t_0, 0) = \tilde{w}_1(\theta) = \frac{1}{4\pi}. \quad (30.3)$$

Если нас интересует условное среднее значение какой-либо функции  $f(\theta)$ , то его можно получить при помощи (30.2), разложив  $f(\theta)$  по полиномам Лежандра и воспользовавшись их

<sup>1)</sup> См. [17], а также [7], § 57.



взаимной ортогональностью:

$$\int_0^{\pi} P_m(\cos \theta) P_n(\cos \theta) \sin \theta d\theta = \frac{2\delta_{mn}}{2n+1}.$$

В некоторых случаях удобнее, не решая уравнения (30.1), получить с его помощью дифференциальное уравнение для самого среднего  $\overline{f(\theta)}$ . Пусть, например,  $f(\theta) = P_1(\theta) = \cos \theta$ . Умножим (30.1) на  $\cos \theta$  и проинтегрируем по  $\theta$  от 0 до  $\pi$ . Учитывая, что

$$\overline{\cos \theta} = \int_0^{\pi} \cos \theta \cdot \tilde{v} \sin \theta d\theta,$$

получаем

$$\frac{d \overline{\cos \theta}}{dt} = -B \overline{\cos \theta}.$$

Решение этого уравнения с начальным условием  $\overline{\cos \theta} = 1$  при  $t = 0$  есть

$$\overline{\cos \theta} = e^{-Bt}. \quad (30.4)$$

При  $t \rightarrow \infty$  это дает  $\overline{\cos \theta} = 0$ , т. е., в соответствии с (30.3), все ориентации становятся равновероятными. Для малых  $t$  имеем

$$\overline{\cos \theta} = 1 - \overline{\theta^2}/2 = 1 - Bt,$$

т. е.

$$\overline{\theta^2} = 2Bt, \quad (30.5)$$

как и для двумерного брауновского движения на плоскости [см. (24.11)].

Если помимо случайных вращающих моментов имеется еще некоторый детерминированный момент, ориентирующий ось частицы в плоскостях  $\varphi = \text{const}$ , то стационарное распределение уже не будет равномерным. Пусть, например, частица обладает дипольным моментом  $p$  и находится в электрическом поле  $E$ . Вращающий момент будет

$$M(\theta) = -pE \sin \theta = -\partial U / \partial \theta,$$

где  $U = -pE \cos \theta$  — энергия во внешнем поле. Систематическая угловая скорость, если пренебречь инерцией частицы, равна  $\dot{\theta} = M/h$ , где  $h = 8\pi\eta a^3$  — вращательный коэффициент вязкого трения. Таким образом, считая по-прежнему, что не зависящие от  $B$  слагаемые в коэффициентах  $A_i$  уравнения Эйнштейна — Фоккера совпадают с правыми частями динамических уравнений для соответствующих скоростей, мы получим теперь

для  $A_\theta$  вместо (27.14) выражение

$$A_\theta = \frac{M}{h} + \frac{B}{2} \operatorname{ctg} \theta.$$

Уравнение (27.17) заменится при этом на

$$\sin \theta \frac{\partial \tilde{v}}{\partial t} = - \frac{\partial}{\partial \theta} \left( \frac{M}{h} \sin \theta \cdot \tilde{v} \right) + \frac{B}{2} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial \tilde{v}}{\partial \theta} \right),$$

а уравнение для стационарной вероятности  $\tilde{w}_1(\theta)$  будет

$$\frac{d}{d\theta} \left[ \sin \theta \left( \frac{B}{2} \frac{d\tilde{w}_1}{d\theta} - \frac{M}{h} \tilde{w}_1 \right) \right] = 0.$$

После однократного интегрирования и учета того, что постоянная интегрирования равна нулю, получим

$$\frac{B}{2} \frac{d\tilde{w}_1}{d\theta} - \frac{M}{h} \tilde{w}_1 = \frac{B}{2} \frac{d\tilde{w}_1}{d\theta} + \frac{1}{h} \frac{dU}{d\theta} \tilde{w}_1 = 0.$$

Отсюда

$$\tilde{w}_1(\theta) = C e^{-2U/hB},$$

причем  $C$  определяется из условия нормировки

$$2\pi \int_0^\pi \tilde{w}_1(\theta) \sin \theta d\theta = 1.$$

Но полученный результат должен совпадать с больцмановским распределением, а значит,  $hB/2 = kT$ , т. е. (см. § 26)

$$B = 2kT/h.$$

Казалось бы, вращательное брауновское движение не имеет никакого отношения к вопросам, интересующим радиофизику. Но предыдущие задачи уже не раз подтверждали, что одна и та же вероятностная схема может охватывать очень разнородные явления. В данном случае мы снова сталкиваемся с примером такого же рода. При определенных ограничениях аналогичной вращательному брауновскому движению оказывается задача о распространении луча в среде с плавными случайными неоднородностями. Говоря о плавности или медленности изменений свойств среды, мы имеем в виду применимость геометрической оптики (или акустики). Для достаточно коротких радиоволн соответствующие условия могут в известной степени выполняться и в тропосфере, и в ионосфере, и в солнечной короне, так что лучевая трактовка случайной рефракции оказывается допустимой. Самое же явление флуктуаций направления распространения луча представляет для радиофизики непосредственный интерес.

Свойства среды могут быть охарактеризованы показателем преломления  $n(x, y, z)$  или же фазовой скоростью волны  $v(x, y, z)$ , изменение которых от точки к точке и описывает неоднородности среды. Утверждение, что эти неоднородности случайны, означает, что  $n$  или  $v$  — случайные функции точки, т. е. не одного параметра  $t$ , как это было всюду ранее, а трех параметров  $x, y, z$ . Это пример так называемого *случайного* поля. В части II этой книги случайные поля будут рассмотрены более подробно, причем не только в связи с задачей о флуктуациях рефракции.

Подобно тому как одной из основных характеристик случайного процесса  $\xi(t)$  является его смешанный момент  $\overline{\xi(t)\xi(t')}$ , так и для случайного поля  $\xi(x, y, z)$ , или, короче,  $\xi(\mathbf{r})$ , момент  $\overline{\xi(\mathbf{r})\xi(\mathbf{r'})}$  чрезвычайно важен, так как он дает меру статистической связи между значениями  $\xi(\mathbf{r})$  в двух точках пространства. Предполагая, что  $\bar{\xi} = 0$ , мы будем далее отождествлять этот смешанный момент с функцией корреляции  $\psi_{\xi}(\mathbf{r}, \mathbf{r'})$ . Поле называется *однородным*, если функция корреляции зависит только от разностей координат этих точек  $\mathbf{R} = \mathbf{r} - \mathbf{r'}$ , т. е. перенос начала отсчета не влияет на  $\psi_{\xi}(\mathbf{R}) = \overline{\xi(\mathbf{r})\xi(\mathbf{r'})}$ . Если же, сверх того,  $\psi_{\xi}$  зависит только от расстояния  $R = |\mathbf{r} - \mathbf{r'}|$ , но не от направления  $\mathbf{R}$ , то поле называется статистически *изотропным*.

Мы предположим, что рассматриваемое случайное поле  $v(\mathbf{r})$  однородно и изотропно, или, иначе говоря, что в среде нет статистически выделенных положений и направлений. Это допущение играет по отношению к случайным поворотам луча при его преломлении в неоднородностях среды такую же роль, как и гипотеза об однородности и изотропности случайных толчков при брауновском движении частицы.

Пространственная функция корреляции однородного и изотропного поля  $\psi_{\xi}(R)$  зависит лишь от одного аргумента, как это имеет место и для временной функции корреляции  $\psi_{\xi}(\tau)$  стационарного случайного процесса. Во многих задачах можно указать для  $\psi_{\xi}(\tau)$  некоторый характерный временной интервал  $\theta$ , в пределах которого корреляция заметна, а при сдвигах  $\tau > \theta$  достаточно быстро убывает. В таких случаях  $\theta$  называют «временем корреляции». В пространственных задачах аналогичную роль играет «радиус корреляции»  $a$ ; при увеличении  $R$  сверх  $a$  функция корреляции  $\psi_{\xi}(R)$  стремится к нулю. Очевидно, в интересующей нас задаче, если понимать под  $\xi$  случайный показатель преломления среды  $n$  (или случайную фазовую скорость волны  $v$ ), то  $a$  характеризует средний размер неоднородностей среды.

Применимость геометрического приближения предполагает, что выполнено не только условие  $a \gg \lambda$  ( $\lambda$  — длина волны), но и условие малости дифракционных эффектов  $z\lambda \ll a^2$ , где  $z$  — толщина слоя, пройденного лучом в неоднородной среде. Последнее условие означает, что даже при наиболее удаленных от точки наблюдения неоднородностях размер зоны Френеля  $\sim \sqrt{z\lambda}$  гораздо меньше размера неоднородностей  $a$ .

По отношению к форме луча  $a$  характеризует протяженность тех его участков, на которых луч испытывает «однократный» поворот. Вследствие этих случайных извилин, обусловленных отдельными неоднородностями, и набегает интересующая нас случайная интегральная рефракция — угловое отклонение  $\theta$  направления луча от оси  $z$ , т. е. от его начального направления  $\theta = 0$ , при входе в неоднородную среду.

Задача о рефракции луча на случайных неоднородностях была тщательно проанализирована в работе [18], в которой было показано, что угол  $\theta$  можно рассматривать как функцию марковского типа при следующих условиях.

1. На всем пути луча должно быть  $\theta \ll \pi/2$  (так называемое *маломуговое приближение*)<sup>1)</sup>, что позволяет не различать в первом приближении толщину  $z$  пройденного слоя среды и полную длину луча  $l$  в этом слое. Иначе это условие выражается требованием, чтобы средний квадрат поперечного смещения луча  $\rho$  от оси  $z$  был мал по сравнению с  $z^2$ :

$$\overline{\rho^2} \ll z^2. \quad (30.6)$$

2. На своем пути в слое толщины  $z$  луч должен испытать очень много «однократных» поворотов, т. е. должен встретить очень много неоднородностей:

$$z \gg a.$$

Неравенство должно быть настолько сильным, чтобы можно было разбить полную толщину  $z$  на отрезки  $\Delta z \ll z$ , причем/уже на каждом таком отрезке луч встречает очень много неоднородностей ( $\Delta z \gg a$ ), т. е. испытывает много взаимно некоррелированных «толчков» из-за случайных градиентов показателя преломления. Вместе с тем средний квадрат результирующего отклонения луча  $\overline{(\Delta\theta)^2}$  на отрезке  $\Delta z$  еще мал — по порядку величины не ниже  $\Delta z$ . Это совершенно аналогично условиям марковости при блужданиях брауновской частицы: за элемент времени  $\Delta t$  частица испытывает очень много независимых толчков, но средний квадрат смещения  $\overline{(\Delta s)^2}$  имеет порядок малости не

<sup>1)</sup> Необходимость этого условия не была замечена в первых работах по случайной рефракции.

ниже, чем  $\Delta t$ . В этом макроскопическом смысле промежуток  $\Delta t$  бесконечно мал.

При этих условиях<sup>1)</sup> существует вероятность перехода  $\bar{v}(\theta|z, 0)$ , удовлетворяющая тому же уравнению (30.1), что и вероятность поворота на угол  $\theta$  оси сферической частицы, подвергающейся действию случайных однородных и изотропных вращающих моментов. То, что вместо времени  $t$  теперь независимой переменной является толщина  $z$  пройденного лучом слоя среды, конечно несущественно. Принципиальное отличие состоит в том, что для частицы угол поворота  $\theta$  не лимитирован, а для луча уравнение (30.1) пригодно лишь при малых  $\theta$ , когда вообще можно заменить  $\sin \theta$  на  $\theta$ . Соответственно вместо (30.4) мы должны считать, что  $\overline{\cos \theta} = 1$ , а для среднего квадрата  $\theta$  можно пользоваться лишь формулой (30.5), т. е.

$$\overline{\theta^2} = 2Bz \ll 1.$$

При этом же ограничении средний квадрат поперечного смещения луча равен

$$\overline{\rho^2} = \frac{2}{3} Bz^3 = \frac{1}{3} \overline{\theta^2} z^2,$$

откуда и вытекает (30.6):

### § 31. Скачкообразные марковские процессы. Уравнение Колмогорова — Феллера

Уравнение Эйнштейна — Фоккера выведено для марковских процессов  $\xi(t)$ , непрерывных не только в смысле непрерывности возможных значений  $x$  случайной функции  $\xi(t)$ , но и в смысле вероятностной непрерывности изменения  $\xi(t)$  во времени («почти все» реализации  $\xi(t)$  непрерывны по  $t$  в обычном смысле). При отказе от этого последнего ограничения, т. е. от условия (26.3), мы получаем возможность рассматривать марковские процессы, у которых изменения состояния происходят *мгновенными скачками*. Общим будет при этом *смешанный* случай, когда состояние может меняться и непрерывно, и скачками, но сначала мы остановимся на чисто скачкообразных переходах. Пример такого процесса уже был упомянут ранее, когда речь шла о невыполнении условия (26.3) при соударениях молекул газа: скорость молекулы меняется при соударении скачком, но множество возможных значений скорости после удара непрерывно. Кроме соударений микрочастиц (молекул, атомов, электронов и т. д.)

<sup>1)</sup> В § 36 мы еще вернемся к этим условиям и выясним, откуда они вытекают.

в эту схему укладываются и квантовые переходы, и некоторые импульсные процессы.

Рассмотрим первоначально следующий простой пример одномерного пуассоновского скачкообразного процесса  $\xi(t)$ . Пусть случайные моменты времени  $t_v$  образуют пуассоновский поток событий, так что вероятность наступления  $n$  событий в интервале  $(0, t)$  дается законом Пуассона

$$P(n) = \frac{(n_1 t)^n}{n!} e^{-n_1 t}, \quad (31.1)$$

где  $n_1 t$  — среднее число событий в  $(0, t)$ ,  $n_1$  — среднее их число в единицу времени.

В момент  $t_v$  процесс  $\xi(t)$  скачком попадает в интервал  $(x_v, x_v + dx_v)^{1)}$  с вероятностью  $\varphi(x_v | x_{v-1}) dx_v$ , зависящей от предшествующего значения  $x_{v-1}$ , принятого в момент  $t_{v-1}$ . Таким образом, если задано начальное значение  $\xi(0) = x_0$ , то вероятность цепочки  $n$  значений  $x_1, \dots, x_n$ , принятых в последовательные моменты скачков  $t_1, \dots, t_n$  ( $t_0 = 0 < t_1 < \dots < t_n$ ), равна

$$\prod_{v=1}^n \varphi(x_v | x_{v-1}) dx_v.$$

Найдем для  $\xi(t)$  плотность условной вероятности  $v(t, x | 0, x_0)$ , которую можно записать и в виде  $v(x | t, x_0)$ , поскольку заранее очевидно, что в силу независимости  $n_1$  и  $\varphi(x_v | x_{v-1})$  от времени процесс однороден по  $t$ . Предположим сначала, что в интервале  $(0, t)$  произошло фиксированное число  $n$  скачков, и найдем плотность условной вероятности  $v(x | t, x_0, n)$  при этом дополнительном условии. Ясно, что

$$v(x | t, x_0) = \sum_{n=0}^{\infty} v(x | t, x_0, n).$$

Так как переход  $x_0 \rightarrow x_n = x$  осуществляется при любых промежуточных значениях  $x_1, \dots, x_{n-1}$ , имеем

$$v(x | t, x_0, n) = P(n) p_n(x | x_0),$$

где

$$p_n(x | x_0) = \int \varphi(x | x_{n-1}) \dots \varphi(x_1 | x_0) dx_1 \dots dx_{n-1}. \quad (31.2)$$

Заметим, что, согласно (31.2),

$$p_n(x | x_0) = \int \varphi(x | x_{n-1}) p_{n-1}(x_{n-1} | x_0) dx_{n-1}. \quad (31.3)$$

<sup>1)</sup>  $v$  — номер скачка, множество же возможных значений  $x_v$  при всяком  $v$  непрерывно.

В частном случае  $n = 0$ , т. е. при отсутствии скачков в  $(0, t)$ , значение  $x$  должно быть равно  $x_0$ , так что  $p_0(x|x_0) = \delta(x - x_0)$  и соответственно

$$v(x|t, x_0, 0) = P(0) \delta(x - x_0).$$

Суммируя теперь  $v(x|t, x_0, n)$  по всем взаимно исключающим частным случаям  $n = 0, 1, 2, \dots$ , мы получаем плотность условной вероятности  $v(x|t, x_0)$ , уже не связанную с фиксированным  $n$ :

$$\begin{aligned} v(x|t, x_0) &= \sum_{n=0}^{\infty} P(n) p_n(x|x_0) = \\ &= P(0) \delta(x - x_0) + \sum_{n=1}^{\infty} P(n) p_n(x|x_0), \end{aligned} \quad (31.4)$$

или, после подстановки  $P(n)$  из (31.1),

$$v(x|t, x_0) = e^{-n_1 t} \left\{ \delta(x - x_0) + \sum_{n=1}^{\infty} \frac{(n_1 t)^n}{n!} p_n(x|x_0) \right\} \quad (31.5)$$

В случае очень короткого интервала времени, заменяя  $t$  на  $\Delta t$  и ограничиваясь членами не выше первого порядка относительно  $\Delta t$ , т. е. пренебрегая многократными скачками ( $n = 2, 3, \dots$ ), мы получаем отсюда для малых промежутков времени

$$v(x|\Delta t, x_0) \approx (1 - n_1 \Delta t) \delta(x - x_0) + n_1 \Delta t \varphi(x|x_0). \quad (31.6)$$

Легко убедиться, что плотность условной вероятности (31.5) удовлетворяет уравнению Смолуховского (см. задачу 17), т. е. является плотностью *вероятности перехода*, определяющей *марковский* процесс  $\xi(t)$ . Кроме того, нетрудно получить для  $v(x|t, x_0)$  интегро-дифференциальное уравнение, решение которого при начальном условии

$$v(x|0, x_0) = \delta(x - x_0) \quad (31.7)$$

дается выражением (31.5). Действительно, дифференцируя (31.5) по  $t$ , получаем

$$\frac{\partial v}{\partial t} = -n_1 v + \sum_{n=1}^{\infty} \frac{(n_1 t)^{n-1}}{(n-1)!} n_1 e^{-n_1 t} p_n(x|x_0),$$

откуда в силу (31.3) и (31.4) следует, что

$$\frac{\partial v(x|t, x_0)}{\partial t} = -n_1 \left[ v(x|t, x_0) - \int \varphi(x|y) v(y|t, x_0) dy \right]. \quad (31.8)$$

Уравнение (31.8) представляет собой очень частный случай уравнения Колмогорова — Феллера — интегро-дифференциаль-

ного уравнения, которому удовлетворяет вероятность перехода марковского случайного процесса с чисто разрывными изменениями состояния (скачками). Это общее уравнение по-прежнему выводится из фундаментального уравнения Смолуховского, но при обобщенном по сравнению с (31.6) выражении для вероятности перехода за малое время  $\Delta t$ . А именно, предполагается, что на интервале  $(t, t + \Delta t)$  процесс  $\xi(t)$  либо переходит (если скачок произошел) от значения  $x$  в момент  $t$  в интервал  $(z, z + dz)$  с условной вероятностью  $\varphi(z|t, x)dz$ , либо сохраняет значение  $x$  (если скачка не было), т. е. в этом последнем случае вероятность попадания в  $(z, z + dz)$  есть  $\delta(z - x)dz$ . В свою очередь пусть скачок в течение малого времени  $\Delta t$  наступает с вероятностью  $a(t, x)\Delta t$ . Следовательно, с точностью до первого порядка относительно  $\Delta t$

$$v(t + \Delta t, z|t, x) \approx [1 - a(t, x)\Delta t]\delta(z - x) + a(t, x)\Delta t \cdot \varphi(z|t, x). \quad (31.9)$$

В отличие от (31.6), временная плотность вероятности скачка  $a(t, x)$  равна здесь не постоянной величине  $n_1$ , а зависит и от начального момента  $t$ , и от значения  $x$  в этот момент. Плотность условной вероятности  $\varphi$  значений  $z$ , принимаемых в результате скачка, тоже зависит не только от предшествующего значения  $x$ , но и от времени  $t$ , так что процесс  $\xi(t)$  уже нестационарен.

В предположении (31.9) и при некоторых дополнительных требованиях (непрерывности  $a$  и  $\varphi$  как функций  $t$  и ограниченности  $a$  на любом конечном интервале  $t$ ) из уравнения Смолуховского вытекает следующее интегро-дифференциальное уравнение для плотности вероятности перехода [17]:

$$\frac{\partial v(t, x|t_0, x_0)}{\partial t} = -a(t, x)v(t, x|t_0, x_0) + \int a(t, y)\varphi(x|t, y)v(t, y|t_0, x_0)dy. \quad (31.10)$$

Это уравнение для  $v$  как функции  $t$  и  $x$  представляет собой аналог уравнения Эйнштейна — Фоккера (26.4) для марковских процессов с непрерывно меняющимся состоянием. В тех же предположениях и таким же путем получается и другое уравнение, аналогичное (26.7), т. е. уравнение для  $v(t, x|t_0, x_0)$  как функций  $t_0$  и  $x_0$ :

$$\frac{\partial v(t, x|t_0, x_0)}{\partial t_0} = a(t_0, x_0) \left[ v(t, x|t_0, x_0) - \int \varphi(y|t_0, x_0)v(t, x|t_0, y)dy \right]. \quad (31.11)$$

При начальном условии

$$v(t_0, x|t_0, x_0) = \delta(x - x_0) \quad (31.12)$$



оба уравнения имеют *единственное* решение, причем *одно и то же* для обоих.

В каком соотношении находится рассматриваемая схема скачкообразного изменения состояний при непрерывном их множестве с исследованным в § 23 случаем *дискретных* возможных состояний, когда из-за самой этой дискретности состояние может меняться только скачками, так сказать «поневоле»?

При дискретных состояниях  $x_k$  плотности условных вероятностей  $v(t + \Delta t, z | t, x_i)$  и  $\varphi(z | t, x_i)$  можно записать через дельта-функции  $\delta(z - x_k)$ :

$$v(t + \Delta t, z | t, x_i) = \sum_k p(t + \Delta t, x_k | t, x_i) \delta(z - x_k),$$

$$\varphi(z | t, x_i) = \sum_k \pi(x_k | t, x_i) \delta(z - x_k),$$

где  $p$  и  $\pi$  — интегральные условные вероятности состояния  $x_k$  ( $p$  — вероятность перехода, не зависящая от того, произошел скачок или нет,  $\pi$  — вероятность при условии, что скачок имел место). Подстановка этих выражений в (31.9) дает

$$\begin{aligned} \sum_k p(t + \Delta t, x_k | t, x_i) \delta(z - x_k) \approx [1 - a(t, x_i) \Delta t] \delta(z - x_i) + \\ + a(t, x_i) \Delta t \sum_k \pi(x_k | t, x_i) \delta(z - x_k). \end{aligned}$$

Интегрируя это равенство по  $z$  (по  $\pm \varepsilon$ -окрестности состояния  $x_k$ ) и учитывая при этом, что

$$\int_{x_k - \varepsilon}^{x_k + \varepsilon} \delta(z - x_i) dz = \delta_{ik},$$

находим следующее выражение для вероятности перехода за малое время  $\Delta t$ :

$$\begin{aligned} p(t + \Delta t, x_k | t, x_i) \approx [1 - a(t, x_i) \Delta t] \delta_{ik} + \\ + a(t, x_i) \Delta t \cdot \pi(x_k | t, x_i) = \delta_{ik} + a(t, x_i) [\pi(x_k | t, x_i) - \delta_{ik}] \Delta t. \end{aligned}$$

Но это выражение имеет вид (23.2):

$$p(t + \Delta t, x_k | t, x_i) \approx \delta_{ik} + A_{ik}(t) \Delta t,$$

т. е. именно тот вид, в предположении которого выводятся уравнения Колмогорова (23.8) для марковских процессов с дискретными возможными состояниями. При этом

$$A_{ik}(t) = a(t, x_i) [\pi(x_k | t, x_i) - \delta_{ik}],$$

чем обеспечивается выполнение условия (23.7):

$$\sum_k A_{ik}(t) = 0.$$

Таким образом, уравнения Колмогорова — Феллера (31.10) и (31.11) для скачкообразных марковских процессов охватывают случаи как непрерывных, так и дискретных возможных состояний. Целесообразно поэтому записывать их для интегральных вероятностей перехода  $V(t, x|t_0, x_0)$ , пользуясь интегралом Стильбеса:

$$\begin{aligned} \frac{\partial V(t, x|t_0, x_0)}{\partial t} = & - \int_{-\infty}^x a(t, y) d_y V(t, y|t_0, x_0) + \\ & + \int_{-\infty}^{+\infty} a(t, y) \pi(x|t, y) d_y V(t, y|t_0, x_0), \quad (31.13) \end{aligned}$$

$$\begin{aligned} \frac{\partial V(t, x|t_0, x_0)}{\partial t_0} = & a(t_0, x_0) \left[ V(t, x|t_0, x_0) - \right. \\ & \left. - \int_{-\infty}^{+\infty} V(t, x|t_0, y) d_y \pi(y|t_0, x_0) \right]. \quad (31.14) \end{aligned}$$

В таком виде эти уравнения и были получены в работе Феллера [19].

Уравнение Колмогорова — Феллера (31.10) можно записать в более лаконичной форме, а именно в виде классического *уравнения Больцмана*, основного уравнения кинетической теории газов. Введем плотность вероятностей скачка в момент времени  $t$  из состояния  $y$  в состояние  $x$ <sup>1)</sup>:

$$u(x|t, y) = a(t, y) \varphi(x|t, y). \quad (31.51)$$

Нетрудно установить справедливость этого равенства. Действительно,  $a(t, y)dt$  — вероятность скачка из  $y$  *куда-нибудь* [т.е. вероятность наличия скачка в  $(t, t+dt)$ ], а  $\varphi(x|t, y)dx$  — условная вероятность того, что *при наличии скачка* процесс перешел в  $(x, x+dx)$ . Произведение  $a$  и  $\varphi$  представляет собой условную плотность вероятностей  $u$ , как она определена выше. Интегрируя (31.15) по всем  $x$ , учитывая при этом, что  $\int \varphi(x|t, y) dx = 1$ , и переобозначая переменные ( $x \rightleftharpoons y$ ), получаем

$$\int u(y|t, x) dy = a(t, x). \quad (31.16)$$

<sup>1)</sup> Разумеется, это лишь краткий способ выражаться. Более точно надо говорить о  $u(x|t, y)dx dt$  как о вероятности попадания в результате скачка в течение времени  $(t, t+dt)$  из состояния  $y$  в интервал  $(x, x+dx)$ .

Если подставить (31.16) в первый член правой части (31.10), а (31.15) — во второй, то уравнение (31.10) примет вид

$$\frac{\partial v(t, x | t_0, x_0)}{\partial t} = -v(t, x | t_0, x_0) \int u(y | t, x) dy + \\ + \int v(t, y | t_0, x_0) u(x | t, y) dy. \quad (31.17)$$

Это и есть уравнение Больцмана (для одномерного процесса). Из него особенно отчетливо видно, что это, по существу, уравнение баланса: скорость изменения  $v$  (например, концентрации частиц в точке  $x$ ) равна разности двух обусловленных скачками ежесекундных потоков: из  $x$  в какие-либо другие состояния и из всех других состояний в  $x$ .

Как уже было сказано, наиболее общим является смешанный случай суперпозиции дискретного и непрерывного «спектров» возможных значений процесса  $\xi(t)$ , когда состояние может меняться как скачками, так и непрерывно. Интегро-дифференциальные уравнения Колмогорова — Феллера для этого общего случая таковы<sup>1)</sup>:

$$\frac{\partial v(t, x | t_0, x_0)}{\partial t} = -\frac{\partial A(x, t) v}{\partial x} + \frac{1}{2} \frac{\partial^2 B(x, t) v}{\partial x^2} - a(t, x) v + \\ + \int_{-\infty}^{+\infty} a(t, y) v(t, y | t_0, x_0) \frac{\partial \pi(x | t, y)}{\partial x} dy, \quad (31.18)$$

$$\frac{\partial V(t, x | t_0, x_0)}{\partial t_0} = -A(x_0, t_0) \frac{\partial V}{\partial x_0} + \frac{B(x_0, t_0)}{2} \frac{\partial^2 V}{\partial x_0^2} + \\ + a(t_0, x_0) \left[ V(t, x | t_0, x_0) - \right. \\ \left. - \int_{-\infty}^{+\infty} V(t, x | t_0, x_0) d_y \pi(y | t_0, x_0) \right], \quad (31.19)$$

где  $A(x, t)$  и  $B(x, t)$  — коэффициенты сноса и диффузии, определенные через *непрерывную часть*  $V$  так же, как это было

<sup>1)</sup> Колмогоров независимо от Феллера указал на существование класса марковских процессов со скачками и в работе [3] привел без вывода уравнения (31.13) и (31.18). Кстати сказать, он предлагал называть процессы без вероятностного последования «стохастически определенными», но этот термин не прижился и установилось название «марковские процессы». В работах [19, 20] Феллер дал анализ условий, при которых получаются интегро-дифференциальные уравнения, а также исследовал вопросы о существовании и единственности решения задачи с начальным условием (31.12).

сделано выше при выводе уравнения Эйнштейна — Фоккера [см. (26.1) и (26.2)]. Уравнение (31.19) вытекает из уравнения Смолуховского при уже указанных выше предположениях относительно  $A$ ,  $B$ ,  $a$  и  $\pi$ , а для вывода (31.18) достаточно допустить, сверх того, существование плотности  $\varphi(x|t, y) = \partial\pi(x|t, y)/\partial x$  и дифференцируемость  $a(t, y)$  и  $\pi(x|t, y)$  по  $t$  и  $y$ . Тогда существует плотность вероятности перехода  $v = \partial V/\partial x$ , для которой и написано уравнение (31.18).

Если вероятность скачка равна нулю ( $a = 0$ ), то мы получаем из (31.18) и (31.19) уравнение Эйнштейна — Фоккера (26.4) и сопряженное уравнение (26.7), предполагающие, конечно, непрерывность множества возможных значений. Если  $a \neq 0$ , но равны нулю коэффициенты сноса и диффузии ( $A = 0$ ,  $B = 0$ ), то мы возвращаемся к уравнениям Колмогорова — Феллера (31.13) и (31.14) для чисто скачкообразных марковских процессов, в том числе для процессов с дискретными возможными значениями.

Вряд ли надо добавлять, что все уравнения легко обобщаются на многомерные марковские процессы  $\xi(t) = \{\xi_1(t), \dots, \xi_n(t)\}$ .

### § 32. Задача о первом достижении границ

Эта задача может быть поставлена и не для марковских процессов, но только для них удается получить ее точное решение в общем виде, что и оправдывает включение этого параграфа в главу о марковских процессах. Даже в частном случае стационарного *нормального* процесса  $\xi(t)$ , для которого, казалось бы, ответ на любой вопрос должен каким-то образом выражаться через функцию корреляции  $\psi(t_1, t_2)$ , в задаче о достижении границ пока получены лишь некоторые частные результаты [21, 22].

Задача ставится следующим образом. Пусть случайный процесс  $\xi(t)$  принял в начальный момент  $t_0 = 0$  значение  $x_0$ . Какова статистика (функции распределения, моменты) *случайного времени*  $t$  первого достижения либо заданной одной границы  $a < x_0$  или  $b > x_0$ , либо достижения одной из них, если заданы обе? Речь идет, таким образом, о том, чтобы, располагая статистикой  $\xi(t)$  «по оси  $x$ » (распределениями  $\omega_n$  значений самой функции), получить статистику «по оси  $t$ ». Задачи такого типа принадлежат к числу наиболее интересных в теории случайных процессов. К ним относятся и вопросы, касающиеся статистики пересечений заданного уровня или «выбросов» случайных процессов [23].

Теория достижения границ *марковским* процессом  $\xi(t)$  впервые была развита Л. С. Понтрягиным, А. А. Андроновым и

А. А. Виттом [24]<sup>1)</sup>. Решение этой задачи находит разнообразные применения. В случае брауновского движения примером может служить опыт Бриллюэна, в котором брауновская частица *прилипает* к стенке сосуда, т. е. речь идет именно о *первом* достижении частицей поглощающей границы [7]. Из радиофизических задач можно указать на вопрос о времени достижения сепаратрисы при флуктуациях в нелинейных динамических системах, в частности о времени перехода через потенциальный барьер в консервативных системах и о времени самопроизвольной смены автоколебательных режимов или срыва автоколебаний в автогенераторах. Сюда же относится задача о срыве слежения в авторегулируемых следящих системах и др.

Следуя [24], мы рассмотрим одномерный марковский процесс<sup>2)</sup>, причем однородный по времени, т. е. с плотностью вероятности перехода  $v(x|t-t_0, x_0)$  и с не зависящими от  $t$  коэффициентами  $A(x)$  и  $B(x)$  в уравнении Эйнштейна — Фоккера. Плотность вероятности перехода  $v$  и одномерная плотность вероятности состояния  $w_1(x)$  (если она существует) определяют всю статистику марковского процесса «по  $x$ ». Но нас будет интересовать другое распределение, отличное от  $w_n$ , а именно вероятность *времени*  $t$  достижения одной из границ области изменения  $x$ :

$$\Phi(t|x) = P \left\{ \begin{array}{l} x=a \text{ или } x=b \text{ по} \\ \text{крайней мере один раз} \\ \text{за время } t \end{array} \middle| \begin{array}{l} \text{если при } t_0=0 \\ \text{имело место} \\ \text{состояние } x \end{array} \right\}.$$

Это *интегральная* вероятность (так что  $\Phi(\infty|x) = 1$ ), плотность же вероятности равна  $\partial\Phi/\partial t$ , и, следовательно, среднее время достижения  $a$  или  $b$  равно

$$\bar{t}(x) = \int_0^\infty t \frac{\partial\Phi}{\partial t} dt. \quad (32.1)$$

Нетрудно установить, каковы начальные и граничные условия для  $\Phi(t|x)$ . Так как в начальный момент  $x \neq a$  и  $x \neq b$ , начальное условие будет

$$\Phi(0|x) = 0. \quad a < x < b. \quad (32.2)$$

Но если при  $t = 0$  либо  $x = a$ , либо  $x = b$ , т. е. граница уже достигнута, то для всякого  $t > 0$  будет справедливо граничное

<sup>1)</sup> Эта работа, опубликованная в 1933 г. и отраженная (в уточненном виде) даже в учебнике [7], по-видимому, осталась неизвестной за рубежом, где начало исследований по теории достижения границ марковским процессом связывают с работами С. Райса (1944—1945 гг.) [8], а также Д. Дарлингга и А. Зигерта (1953) [25].

<sup>2)</sup> Теория легко обобщается и на многомерный случай [7].

условие

$$\varphi(t|a) = \varphi(t|b) = 1. \quad (32.3)$$

Теперь надо составить уравнение для  $\varphi(t|x)$  — задача, которая полностью решается именно для марковского процесса при помощи вероятности перехода. Нам нужна при этом не вероятность перехода  $v(y|t, x)dy$  из точки  $x$  при  $t = 0$  в *любой* интервал  $(y, y + dy)$  на оси  $x$ , а вероятность  $\hat{v}(y|t, x)dy$  того же события, но с  $y$ , лежащим *внутри* интервала  $(a, b)$ . Конечно, начальные условия для  $v$  и  $\hat{v}$  одинаковы:

$$\hat{v}(y|0, x) = v(y|0, x) = \delta(y - x)$$

(поэтому при очень малых  $t$ , когда  $v$  и  $\hat{v}$  еще очень острые функции  $y$ , они практически совпадают). Но при любых  $t > 0$ , очевидно,  $\hat{v} \leq v$  и условия нормировки для  $v$  и  $\hat{v}$  тоже различны. В то время как

$$\int_{-\infty}^{+\infty} v(y|t, x) dy = 1,$$

для  $\hat{v}$  имеем

$$\int_a^b \hat{v}(y|t, x) dy = 1 - \varphi(t|x). \quad (32.4)$$

Действительно, пусть  $P_n$  — вероятность прихода к моменту времени  $t$  в какую-либо точку интервала  $(a, b)$  с  $n$  достижениями (касаниями, пересечениями) границ по дороге. Полная система возможных событий отвечает  $n = 0, 1, 2, \dots$ , так что  $\sum_{n=0}^{\infty} P_n = 1$ . Но левая часть (32.4) есть  $P_0$  [вероятность прихода куда-либо в  $(a, b)$  без касания границ], а  $\varphi(t|x)$  — вероятность того же конечного результата, но с достижением границ *по крайней мере один раз*, т. е.

$$\varphi(t|x) = P_1 + P_2 + \dots, \quad (32.5)$$

из чего и следует (32.4).

Пользуясь  $\hat{v}$ , можно записать следующее интегральное уравнение (рис. 25):

$$\varphi(t + \tau|x) = \varphi(\tau|x) + \int_a^b \hat{v}(y|\tau, x) \varphi(t|y) dy, \quad (32.6)$$

т. е. вероятность хотя бы одного достижения  $a$  или  $b$  за время  $t + \tau$  (левая часть) равна вероятности того же события за время  $\tau$  [член  $\varphi(\tau|x)$  в правой части] плюс вероятность того, что за  $\tau$  не было касания границ, но оно произошло в течение

остального времени  $t$ . Именно этот смысл имеет второй член в правой части. Действительно,  $\hat{v}(y|\tau, x)dy$  — вероятность перехода за время  $\tau$  из  $x$  в  $y$  без касания границ, а  $\varphi(t|y)$  — вероятность хотя бы одного касания за время  $t$ , так что  $\hat{v}(y|\tau, x)\varphi(t|y)dy$  — это вероятность хотя бы одного касания в интервале времени  $(0, t + \tau)$  с прохождением через  $(y, y + dy)$  в момент  $\tau$ . Интеграл по всем  $y$  от  $a$  до  $b$  и дает полную вероятность того, что за  $\tau$  касания границ не было, но оно имело место в течение остальной части  $t$  промежутка времени  $t + \tau$ . Заметим, что, согласно (32.6), разность

$$\begin{aligned} \varphi(t + \tau|x) - \varphi(t|x) &= \\ &= \int_a^b \hat{v}(y|\tau, x) \varphi(\tau|y) dy \end{aligned}$$

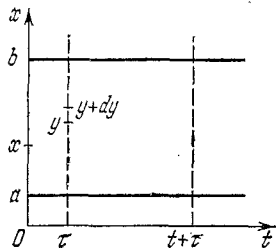


Рис. 25.

есть вероятность достижения границы (хотя бы один раз) в интервале  $(t, t + \tau)$  после «пустого» интервала  $(0, t)$ . Заменяя  $\tau$  на  $dt$ , мы получаем, что  $\frac{\partial \varphi}{\partial t} dt$  представляет собой с точностью до первого порядка относительно  $dt$  вероятность *первого* достижения границы в интервале  $(t, t + dt)$ .

Предельный переход  $\tau \rightarrow 0$  в уравнении (32.6) вполне строго приводит к (сопряженному) уравнению Эйнштейна — Фоккера для  $\varphi(t|x)$  [7]. Но тот же результат можно получить и при некотором физически очевидном упрощении исходного интегрального уравнения (32.6), что и было сделано в работе [24]. Поскольку  $\tau$  мало (а в дальнейшем  $\tau \rightarrow 0$ ), можно сразу же пренебречь в (32.6) членом  $\varphi(\tau|x)$ , отождествить острые функции  $\hat{v}$  и  $v$  и раздвинуть пределы интеграла в  $\pm \infty$ , т. е. исходить из уравнения

$$\varphi(t + \tau|x) = \int_{-\infty}^{+\infty} v(y|\tau, x) \varphi(t|y) dy. \quad (32.7)$$

Разлагая  $\varphi(t|y)$  по степеням  $y - x$ :

$$\begin{aligned} \varphi(t|y) &= \varphi(t|x) + \frac{\partial \varphi(t|x)}{\partial x} (y - x) + \\ &+ \frac{1}{2} \frac{\partial^2 \varphi(t|x)}{\partial x^2} (y - x)^2 + O[(y - x)^3], \end{aligned}$$

подставляя это разложение в (32.7) и деля все уравнение на  $\tau$ , получаем (при тех же предположениях, какие были сделаны в § 26 при выводе уравнения Эйнштейна — Фоккера для вероят-

ности перехода  $v$ ) уравнение

$$\frac{\partial \Phi(t|x)}{\partial t} = A(x) \frac{\partial \Phi}{\partial x} + \frac{B(x)}{2} \frac{\partial^2 \Phi}{\partial x^2}, \quad (32.8)$$

причем  $A(x)$  и  $B(x)$  те же, что и ранее. Конечно, этот упрощенный вывод оставляет некоторую неудовлетворенность, но, как сказано, уравнение (32.8) получается и в результате строгого вывода из (32.6).

Можно либо решать уравнение (32.8) с начальными и граничными условиями (32.2) и (32.3), либо перейти к уравнению для характеристической функции:

$$f_t(u|x) = e^{iut} = \int_0^\infty e^{iut} \frac{\partial \Phi(t|x)}{\partial t} dt,$$

либо составить обыкновенные дифференциальные уравнения для моментов  $\bar{t}^n(x)$ . Сделаем последнее.

Дифференцируя (32.8) по  $t$ , умножая результат на  $t^n$  и интегрируя по  $t$  от 0 до  $\infty$ , получаем

$$A(x) \frac{d\bar{t}^n}{dx} + \frac{B(x)}{2} \frac{d^2 \bar{t}^n}{dx^2} = \int_0^\infty t^n \frac{\partial^2 \Phi}{\partial t^2} dt = t^n \frac{\partial \Phi}{\partial t} \Big|_0^\infty - n \bar{t}^{n-1}.$$

В предположении, что плотность вероятности  $\partial \Phi / \partial t$  ограничена в нуле и достаточно быстро стремится к нулю на бесконечности, получаем

$$\frac{B(x)}{2} \frac{d^2 \bar{t}^n}{dx^2} + A(x) \frac{d\bar{t}^n}{dx} + n \bar{t}^{n-1} = 0. \quad (32.9)$$

В частности, при  $n = 1$  и  $n = 2$

$$\frac{B(x)}{2} \bar{t}'' + A(x) \bar{t}' + 1 = 0, \quad (32.10)$$

$$\frac{B(x)}{2} \bar{t}^{2''} + A(x) \bar{t}^{2'} + 2\bar{t} = 0. \quad (32.11)$$

Очевидно, решения таких уравнений должны быть неотрицательны и должны обращаться в нуль всякий раз, когда начальное состояние  $x$  совпадает с одной из границ:

$$\bar{t}^n(a) = \bar{t}^n(b) = 0, \quad (32.12)$$

Если же имеется только одна граница, скажем  $b > x$ , то на ней по-прежнему  $\bar{t}^n(b) = 0$ , а при  $x \rightarrow -\infty$  естественно требовать  $\bar{t}^n \rightarrow \infty$ .

Рассмотрим некоторые примеры.



Пусть  $A(x) = 0$ , а  $B = \text{const}$ . Уравнение (32.8) принимает тогда вид

$$\dot{\varphi} = \frac{B}{2} \varphi'', \quad (32.13)$$

и его решение при условиях (32.2) и (32.3) есть

$$\varphi(t|x) = 1 + \frac{2}{\sqrt{\pi}} \left[ \int_{(b-x)/\sigma}^{(b-a)/\sigma} e^{-s^2} ds - \int_0^{(x-a)/\sigma} e^{-s^2} ds \right], \quad \sigma = \sqrt{2Bt}. \quad (32.14)$$

Ход  $\varphi(t|x)$  в функции от  $x$  показан на рис. 26, а для различных

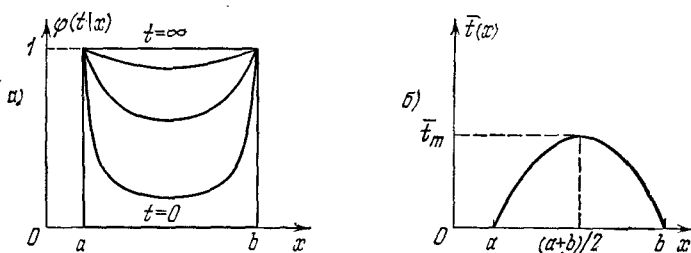


Рис. 26.

значений  $t$ . Для среднего времени достижения границы имеем из (32.10) уравнение

$$\frac{B}{2} \bar{t}'' = -1,$$

решение которого с условиями (32.12) есть

$$\bar{t}(x) = \frac{(b-x)(x-a)}{B} \quad (32.15)$$

(рис. 26, б). Наибольшее время отвечает начальному положению посередине между границами  $x = (a+b)/2$  и равно времени диффузии (с коэффициентом диффузии  $B/2$ ) на половинную длину отрезка  $(a, b)$ :

$$2 \cdot \frac{B}{2} \bar{t}_m = \left( \frac{b-a}{2} \right)^2.$$

Отодвинем теперь нижнюю границу в  $-\infty$ . Тогда решение уравнения (32.13) должно удовлетворять условиям

$$\varphi(t|b) = 1, \quad \varphi(t|-\infty) = 0.$$

Но именно такое решение получается из (32.14) при  $a \rightarrow -\infty$ :

$$\varphi(t|x) = 1 - \frac{2}{\sqrt{\pi}} \int_0^{(b-x)/\sigma} e^{-s^2} ds, \quad \sigma = \sqrt{2Bt}.$$

Соответствующая плотность вероятности равна

$$\frac{\partial \varphi}{\partial t} = \frac{2}{\sqrt{2\pi Bt}} e^{-(b-x)^2/2Bt}.$$

Она обращается в  $\delta(b-x)$  при  $t=0$ , но интеграл (32.1) расходится на верхнем пределе, так как с ростом  $t$  плотность  $\partial\varphi/\partial t$  убывает лишь как  $1/\sqrt{t}$ , т.е. удельный вес больших времен достижения границы слишком велик. Тем самым,  $\bar{t}(b)=0$ , но для всякого  $x < b$  мы получаем  $\bar{t}(x) = \infty$ . То же самое следует и из формулы (32.15) при предельном переходе  $a \rightarrow -\infty$ .

В общем случае переменных коэффициентов  $A(x)$  и  $B(x)$  и при наличии *только одной* границы, (на которой  $\bar{t}=0$ ), если потребовать наиболее медленного роста  $\bar{t}(x)$  при удалении  $x$  от границы, получаем следующие решения уравнения (32.10):

$$\begin{aligned} \bar{t}_a(x) &= 2 \int_a^x e^{-q(z)} dz \int_z^\infty e^{q(y)} \frac{dy}{B(y)} \quad \text{для } a < x, \\ \bar{t}_b(x) &= 2 \int_x^b e^{-q(z)} dz \int_{-\infty}^z e^{q(y)} \frac{dy}{B(y)} \quad \text{для } b > x, \end{aligned}$$

где

$$q(z) = 2 \int_0^z \frac{A(z)}{B(z)} dz.$$

### Задачи

1. Показать, что уравнение Смолуховского следует из определения (15.3) марковского процесса и согласованности конечномерных распределений.

Решение. Согласно (15.3) при  $t_0 < \theta < t$

$$w_3(x_0, t_0; y, \theta; x, t) = w_1(x_0, t_0) v(y, \theta | x_0, t_0) v(x, t | y, \theta),$$

$$w_2(x_0, t_0; x, t) = w_1(x_0, t_0) v(x, t | x_0, t_0).$$

По условию согласованности

$$w_2(x_0, t_0; x, t) = \int w_3(x_0, t_0; y, \theta; x, t) dy.$$

Подставив сюда  $w_2$  и  $w_3$ , получаем уравнение (21.1).

2. Пусть  $\xi_m$  ( $m=1, 2, \dots$ ) — дискретная последовательность независимых случайных величин с возможными значениями  $x_i = i = 0, 1, 2, \dots$  и со-

ответствующими вероятностями  $p_i = P\{\xi_m = i\}$ ,  $\sum_{i=0}^{\infty} p_i = 1$ . Вероятности перехода (совпадающие с вероятностями состояний) не зависят от номера испытания  $m$ , т. е. последовательность однородна:

$$p_{ij} \equiv P(x_j, m+1 | x_i, m) = P\{\xi_{m+1} = j | \xi_m = i\} = P\{\xi_{m+1} = j\} = p_j.$$

Показать, что последовательность величин

$$\eta_m = \xi_1 + \xi_2 + \dots + \xi_m$$

(по определению  $\eta_0 = 0$ ) является марковской, и найти для нее матрицу  $\pi_1$  вероятностей перехода за один шаг.

Решение. Последовательность  $\eta_m$  марковская, поскольку  $\eta_{m+1} = \eta_m + \xi_{m+1}$ , и, следовательно, если задано значение  $\eta_m$ , то распределение  $\eta_{m+1}$  известно. Обозначив через  $y_i = i$  возможные значения  $\eta_m$ , запишем вероятность перехода за один шаг:

$$\begin{aligned} p_{ij} &\equiv P\{y_j, m+1 | y_i, m\} = P\{\eta_{m+1} = j | \eta_m = i\} = \\ &= P\{\eta_m + \xi_{m+1} = j | \eta_m = i\} = P\{\xi_{m+1} = j - i\} = \begin{cases} p_{j-i} & \text{при } j \geq i, \\ 0 & \text{при } j < i. \end{cases} \end{aligned}$$

Таким образом, матрица  $\pi_1$  имеет вид

$$\pi_1 = \begin{pmatrix} p_0 & p_1 & p_2 & p_3 & \dots \\ 0 & p_0 & p_1 & p_2 & \dots \\ 0 & 0 & p_0 & p_1 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}.$$

3. Независимые случайные величины  $\xi_m$  ( $m = 1, 2, \dots$ ) принимают значения  $+1$  или  $-1$  с вероятностями  $p$  и  $q = 1 - p$ . Является ли последовательность величин  $\eta_m = (\xi_m + \xi_{m+1})/2$  марковской?

Решение. Величина  $\eta_m$  может принимать значения  $i = -1, 0, +1$  с вероятностями  $p_i = P\{\eta_m = i\}$ , равными соответственно  $q^2$ ,  $2qp$  и  $p^2$  (одномерное распределение  $\eta_m$ ). Для двух величин  $\eta_m$  и  $\eta_n$  совместная вероятность (двумерное распределение)  $p_{ij} = P\{\eta_m = i, \eta_n = j\}$  будет различна в зависимости от разности  $n - m$ . Если  $|n - m| > 1$ , то  $\eta_m$  и  $\eta_n$  независимы, так что  $p_{ij} = p_i p_j$ . Если  $n = m + 1$ , то в  $\eta_m$  и  $\eta_{m+1}$  входит одна и та же величина  $\xi_m$  и матрица  $p_{ij}$  будет

| $i \backslash j$ | $-1$   | $0$    | $1$    |
|------------------|--------|--------|--------|
| $-1$             | $q^3$  | $pq^2$ | $0$    |
| $0$              | $pq^2$ | $pq$   | $p^2q$ |
| $1$              | $0$    | $p^2q$ | $p^3$  |

Очевидно, такая же матрица получится и при  $n = m + 1$ , когда  $\xi_{m+1}$  входит и в  $\eta_m$  и в  $\eta_{m+1}$ . Наконец, если  $n = m$ , то  $p_{ij} = p_i \delta_{ij}$ .

Для условной вероятности

$$p(n, j | m, i) = P\{\eta_n = j | \eta_m = i\} = \frac{P\{\eta_m = i, \eta_n = j\}}{P\{\eta_m = i\}} = \frac{p_{ij}}{p_i}$$

получаем в трех указанных случаях следующие матрицы:

$$p(n, j | m, i) = \begin{pmatrix} q^2 & 2pq & p^2 \\ q^2 & 2pq & p^2 \\ q^2 & 2pq & p^2 \end{pmatrix} \quad \text{при } |n - m| > 1, \quad (1)$$

$$= \begin{pmatrix} q & p & 0 \\ q/2 & 1/2 & p/2 \\ 0 & q & p \end{pmatrix} \quad \text{при } |n - m| = 1, \quad (2)$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{при } n - m = 0. \quad (3)$$

Если  $\eta_m$  образуют марковскую последовательность, то  $p(n, j | m, i)$  должны быть вероятностями перехода, т. е. удовлетворять уравнению Смолуховского:

$$p(n, k | l, i) = \sum_j p(n, k | m, j) p(m, j | l, i), \quad l < m < n.$$

Однако нетрудно убедиться, что это не имеет места. Достаточно взять случай  $n = m + 1$ ,  $l = m - 1$ , так что  $n - l = 2$  и левая часть уравнения должна быть матрицей (1). Между тем сумма, стоящая в правой части уравнения и представляющая собой квадрат матрицы (2), есть матрица

$$\begin{pmatrix} q\left(q + \frac{p}{2}\right) & p\left(q + \frac{1}{2}\right) & \frac{p^2}{2} \\ \frac{q}{2}\left(q + \frac{1}{2}\right) & pq + \frac{1}{4} & \frac{p}{2}\left(p + \frac{1}{2}\right) \\ \frac{q^2}{2} & q\left(p + \frac{1}{2}\right) & p\left(p + \frac{q}{2}\right) \end{pmatrix}.$$

4. Написать матрицу вероятностей перехода за один шаг для рассмотренного в § 22 примера случайных блужданий (рис. 11), но в случае *поглощающих* стенок.

Решение. Очевидно, при  $1 < i < N$  вероятности перехода  $p_{ik}$  остаются теми же, что и в случае отражающих стенок. Для частицы же, попавшей на стенку ( $i = 1$  или  $i = N$ ), движение останавливается, т. е.

$$p_{1k} = \begin{cases} 1 & (k = 1), \\ 0 & (k \neq 1), \end{cases} \quad p_{Nk} = \begin{cases} 1 & (k = N), \\ 0 & (k \neq N). \end{cases}$$

Таким образом,

$$\pi_1 = (p_{ik}) = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ q & 0 & p & 0 & \dots & 0 & 0 & 0 \\ 0 & q & 0 & p & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & q & 0 & p \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{pmatrix}.$$

5. Доказать, что у нормального стационарного марковского процесса коэффициент корреляции равен

$$K(\tau) = e^{-\alpha |\tau|},$$

где  $\alpha > 0$  (теорема Дуба, см. [26], § 8, а также [27]).

Решение. Согласно (7.13) условная вероятность у нормального процесса с  $\bar{x} = 0$  и  $\bar{x}^2 = 1$  равна

$$v(x_2 | x_1) = \frac{w_2(x_1, x_2)}{w_1 w_2} = \frac{1}{\sqrt{2\pi(1-K^2)}} \exp \left\{ -\frac{(Kx_1 - x_2)^2}{2(1-K^2)} \right\}, \quad (1)$$

где  $K = \overline{x_1 x_2} = K(t_2 - t_1)$  — коэффициент корреляции (процесс по условию стационарен и  $t_2 > t_1$ ). Поскольку процесс марковский,  $v$  должна удовлетворять уравнению Смолуховского:

$$v(x_2 | x_1) = \int_{-\infty}^{+\infty} v(x_2 | y) v(y | x_1) dy.$$

Подставив сюда выражение вида (1) для  $v(x_2 | y)$  с  $K_2 = \overline{x_2 y} = K(t_2 - \theta)$  и для  $v(y | x_1)$  с  $K_1 = \overline{y x_1} = K(\theta - t_1)$  и выполнив интегрирование по  $y$ , получаем

$$\int_{-\infty}^{+\infty} v(x_2 | y) v(y | x_1) dy = \frac{1}{\sqrt{2\pi(1-K_2^2 K_1^2)}} \exp \left\{ -\frac{(K_2 K_1 x_1 - x_2)^2}{2(1-K_2^2 K_1^2)} \right\},$$

что должно быть равно (1). Следовательно,  $K = K_2 K_1$ , или

$$K(t_2 - t_1) = K(t_2 - \theta) K(\theta - t_1), \quad t_2 > \theta > t_1.$$

Для положительных аргументов решение этого функционального уравнения есть

$$K(\tau) = e^{-\alpha\tau}, \quad \tau \geq 0,$$

где  $\alpha$  — любое комплексное число. Так как  $K$  должно удовлетворять условиям  $|K| \leq 1$ , и  $K(-\tau) = K^*(\tau)$ , необходимо, чтобы  $\alpha$  было вещественно и положительно. Отсюда же следует, что при продолжении функции  $K(\tau)$  в область  $\tau < 0$  должно выполняться условие ее четности.

6. Найти связь вероятности перехода  $p_{N,m} = p(N, m | 0, 0)$  из точки 0 в точку  $m$  за  $N$  шагов с биномиальной вероятностью  $P_N(n)$  того, что из общего числа  $N$  независимых шагов сделано вправо  $n$ .

Решение.  $P_N(n)$  — это вероятность перемещения вправо на  $2n - N$  шагов (из любой начальной точки). Если нас интересует смещение вправо на  $m$  шагов, т. е.  $m = 2n - N$ , то вправо должно быть сделано  $n = (m + N)/2$  шагов, что возможно, конечно, только при одинаковой четности  $N$  и  $m$ . Таким образом,

$$p_{N,m} = P_N \left( \frac{m+N}{2} \right), \quad \text{где } m+N \text{ четно.}$$

Нетрудно проверить подстановкой, что это выражение удовлетворяет как уравнению Маркова

$$p_{N,m} = p \cdot p_{N-1,m-1} + q \cdot p_{N-1,m+1},$$

так и условию нормировки  $\sum_{m=-N}^N p_{N,m} = 1$  (сумма по значениям  $m$  той же четности, что и  $N$ ).

7. Найти вероятности перехода  $p(x_k | t, x_i) \equiv p_{ik}(t)$  для однородного по  $t$  марковского процесса с дискретными состояниями, если (постоянные)

коэффициенты  $A_{jk}$  имеют значения

$$A_{j,j+1} = n_1, \quad A_{jj} = -n_1, \quad A_{jk} = 0 \quad \text{при остальных } k,$$

а начальные условия  $p_{ik}(0) = \delta_{ik}$  (Колмогоров [3]).

Ответ.

$$p_{ik}(t) = \begin{cases} \frac{(n_1 t)^{k-i}}{(k-i)!} e^{-n_1 t} & \text{при } k \geq i, \\ 0 & \text{при } k < i. \end{cases}$$

8. Найти стационарные вероятности состояния  $P_k$  для однородного по  $t$  марковского процесса со следующими (постоянными) значениями коэффициентов  $A_{jk}$  ( $j, k = 1, 2, \dots$ ):

$$A_{j,j+1} = \alpha, \quad A_{j,j-1} = \beta > \alpha, \quad A_{11} = -\alpha, \quad A_{jj} = -(\alpha + \beta) \quad \text{при } j > 1, \\ A_{jk} = 0 \quad \text{при остальных } k$$

(Колмогоров [3]).

Решение. Уравнения (23.14) имеют в данном случае вид

$$\beta P_2 - \alpha P_1 = 0 \quad (k=1),$$

$$\beta P_{k+1} - \alpha P_k = \beta P_k - \alpha P_{k-1} \quad (k > 1),$$

т. е.  $\beta P_k = \alpha P_{k-1}$  для всех  $k$ . Следовательно,  $P_k = P_1 (\alpha/\beta)^{k-1}$ , а из условия  $\sum_{k=1}^{\infty} P_k = 1$  следует, что  $P_1 = 1 - \alpha/\beta$ . Окончательно  $P_k = (1 - \alpha/\beta) (\alpha/\beta)^{k-1}$ .

9. Задача ставится так же, как и предыдущая, но со следующими значениями коэффициентов  $A_{jk}$  ( $j, k = 1, 2, \dots$ ):

$$A_{j,j+1} = \alpha, \quad A_{j,j-1} = (j-1)\beta, \quad A_{jj} = -\alpha - (j-1)\beta, \\ A_{jk} = 0 \quad \text{при остальных } k$$

(Колмогоров [3]).

Ответ.

$$P_{k+1} = \frac{e^{-\alpha/\beta}}{k!} \left( \frac{\alpha}{\beta} \right)^k.$$

10. Флуктуации почернения фотоэмульсии [28]. Пусть область эмульсии, соответствующая наименьшему разрешимому точечному объекту (коротко — ячейка), содержит в момент начала экспонирования  $t = 0$   $N$  «непроявленных» зерен. Каждое из них может находиться только в двух состояниях — непроявленном и проявленном. Если к моменту  $t$  остались непроявленными  $m$  зерен ( $m = 1, 2, \dots, N$ ), то предполагается, что вероятность проявления какого-либо одного из них в интервале  $(t, t+dt)$  равна  $A_{m,m-1}(t)dt = m\alpha(t)dt$ , а вероятности проявления двух или более зерен имеют более высокий порядок относительно  $dt$ .

Составить и решить систему уравнений Колмогорова для вероятности перехода  $p(t, m|0, N) \equiv p_m(t)$  ( $m = 1, 2, \dots, N$ ) от  $N$  непроявленных зерен в момент  $t = 0$  к  $m$  таким зернам в момент  $t$ . Найти  $\bar{m}$  и  $D[m]$ .

Решение. При  $m$  непроявленных зернах в момент  $t$  возможен за последующий интервал  $dt$  либо переход к  $m-1$  непроявленным зернам с вероятностью  $m\alpha(t)dt$ , либо сохранение числа непроявленных зерен с вероятностью  $1 - A_{m,m}(t)dt$ , причем  $A_{m,m} + A_{m,m-1} = 0$ , т. е.  $A_{m,m}(t) = -m\alpha(t)$ . Следовательно, при  $m < N$  уравнение для  $p_m(t)$  будет

$$\frac{dp_m}{dt} = A_{m,m}p_m + A_{m+1,m}p_{m+1} = -m\alpha(t)p_m + (m+1)\alpha(t)p_{m+1},$$

а при  $m = N$ , поскольку  $p_{N+1} = 0$ ,

$$\frac{dp_N}{dt} = -N\alpha(t)p_N.$$

Решение этих уравнений при начальном условии  $p_m(0) = \delta_{Nm}$  дается биномиальным законом:

$$p_m(t) = \frac{N!}{(N-m)!m!} p^m(t) q^{N-m}(t),$$

где

$$p(t) = e^{-\rho(t)}, \quad q(t) = 1 - p(t) = 1 - e^{-\rho(t)}, \quad \rho(t) = \int_0^t \alpha(\theta) d\theta.$$

Очевидно,  $p(t)$  — вероятность того, что за время  $t$  зерно останется непроявленным. Таким образом, среднее значение и дисперсия  $m$  таковы:

$$\bar{m} = Np(t) = Ne^{-\rho(t)}, \quad \sigma^2 \equiv D[m] = Np(t)q(t) = Ne^{-\rho(t)}[1 - e^{-\rho(t)}].$$

Процесс нестационарен, и  $D[m]$  зависит от  $\bar{m}$ :

$$\sigma^2 = \bar{m} \left(1 - \frac{\bar{m}}{N}\right) = \bar{n} \left(1 - \frac{\bar{n}}{N}\right),$$

где  $\bar{n} = N - \bar{m}$  — среднее число проявленных зерен.

С ростом  $N$  биномиальное распределение переходит (при фиксированном  $\bar{m}$ ) в пуассоновское. Поэтому в задаче 7, где сразу предполагалось, что  $N = \infty$ , получился закон Пуассона, причем стационарный, так как  $A_{jk}$  там приняты постоянными.

Согласно [28], хорошей аппроксимацией для  $\alpha(t)$  может служить линейная функция времени экспозиции  $t$ :  $\alpha(t) = t/\theta^2$ . Тогда  $\rho(t) = t^2/2\theta^2$  и

$$\bar{n} = N - \bar{m} = N[1 - e^{-\rho(t)}] = N(1 - e^{-t^2/2\theta^2}).$$

По такому же закону нарастает с  $t$  и почернение, пропорциональное  $\bar{n}$ . Следовательно,  $\theta$  — такая длительность экспозиции, при которой скорость роста почернения максимальна. Разумеется,  $\theta$  зависит от освещенности эмульсии. В отличие от почернения, растущего с  $t$  монотонно, фотографический «шум» (т. е.  $\sigma^2$ ) максимален при  $t = \theta\sqrt{\ln 4}$ , достигая здесь значения  $\sigma^2 = N/4$ . При  $t = 0$  и  $t = \infty$   $\sigma^2 = 0$ . «Шум» тем сильнее, чем больше зерен в ячейке, так как  $\sigma^2 \sim N$ .

11. Вывести методом Релея (§ 24) уравнение Эйнштейна — Фоккера для вероятности перехода  $v(t, x) \equiv v(t, x|0, 0)$  в случае, когда частица может совершать за время  $\Delta t$  шаги  $+ka$  и  $-ka$  с вероятностями соответственно  $p_k$  и  $q_k$  ( $k = 1, 2, \dots, n$ , так называемые  $n$ -шаговые блуждания), где, конечно,

$$\sum_{k=1}^n (p_k + q_k) = 1.$$

Решение. Полагая  $t = N\Delta t$ ,  $x = ma$ , получаем для вероятности перехода  $p_{N,m}$  уравнение Маркова:

$$p_{N,m} = \sum_{k=1}^n (p_k p_{N-1, m-k} + q_k p_{N-1, m+k}),$$

или

$$v(t, x) = \sum_{k=1}^n [p_k v(t - \Delta t, x - ka) + q_k v(t - \Delta t, x + ka)].$$

Разложение правой части в ряд по степеням  $a$  дает

$$v(t, x) - v(t - \Delta t, x) = -a \sum_{k=1}^n (p_k - q_k) k \frac{\partial v}{\partial x} + \frac{a^2}{2} \sum_{k=1}^n (p_k + q_k) k^2 \frac{\partial^2 v}{\partial x^2} - \dots$$

Если предположить, что при  $\Delta t \rightarrow 0$  порядок  $a$  и  $\sum_{k=1}^n (p_k - q_k) k$  есть  $\sqrt{\Delta t}$ , то в пределе отличны от нуля будут только члены с  $\partial v / \partial x$  и  $\partial^2 v / \partial x^2$ , т. е. получится диффузионное уравнение

$$\frac{\partial v(t, x)}{\partial t} = -A \frac{\partial v}{\partial x} + \frac{B}{2} \frac{\partial^2 v}{\partial x^2},$$

где

$$A = \lim_{\Delta t \rightarrow 0} \frac{a}{\Delta t} \sum_{k=1}^n (p_k - q_k) k, \quad B = \lim_{\Delta t \rightarrow 0} \frac{a^2}{\Delta t} \sum_{k=1}^n (p_k + q_k) k^2.$$

Пусть, например,

$$p_k = p_1 e^{-(k-1)\alpha}, \quad q_k = q_1 e^{-(k-1)\alpha}$$

( $\alpha > 0$  и  $k = 1, 2, \dots$ , т. е.  $n = \infty$ ). Тогда из условия нормировки получаем

$$\sum_{k=1}^{\infty} (p_k + q_k) = \frac{p_1 + q_1}{1 - e^{-\alpha}} = 1, \quad \text{т. е. } p_1 + q_1 = 1 - e^{-\alpha}.$$

Далее,

$$\sum_{k=1}^{\infty} p_k k = -p_1 e^{\alpha} \frac{\partial}{\partial \alpha} \sum_{k=1}^{\infty} e^{-k\alpha} = \frac{p_1}{(1 - e^{-\alpha})^2}, \quad \sum_{k=1}^{\infty} q_k k = \frac{q_1}{(1 - e^{-\alpha})^2}.$$

Аналогично получаем

$$\sum_{k=1}^{\infty} p_k k^2 = p_1 \frac{1 + e^{-\alpha}}{(1 - e^{-\alpha})^3}, \quad \sum_{k=1}^{\infty} q_k k^2 = q_1 \frac{1 + e^{-\alpha}}{(1 - e^{-\alpha})^3}.$$

В результате

$$A = \frac{A_0}{(1 - e^{-\alpha})^2}, \quad B = \frac{B_0 (1 + e^{-\alpha})}{(1 - e^{-\alpha})^2},$$

где

$$A_0 = \lim_{\Delta t \rightarrow 0} \frac{(p_1 - q_1) a}{\Delta t}, \quad B_0 = \lim_{\Delta t \rightarrow 0} \frac{a^2}{\Delta t}.$$

При  $\alpha \rightarrow \infty$  отличны от нуля только  $p_1$  и  $q_1$ , и мы возвращаемся к случаю одношаговых блужданий, рассмотренному в § 24, причем  $A = A_0$  и  $B = B_0$ . При  $\alpha \rightarrow 0$  вероятности многошаговых скачков возрастают и поэтому ускоряется как систематическое движение, так и диффузия:

$$\bar{s} \approx At \approx \frac{A_0}{\alpha^2} t, \quad \overline{(\Delta s)^2} = 2Bt \approx \frac{4B_0}{\alpha^2} t.$$



12. Рассчитать по формуле (25.9) распределение модуля вектора  $г$ , компоненты которого  $x$  и  $y$  независимы и распределены равномерно в интервале  $(-A, A)$ , т. е.

$$w(x) = \begin{cases} 1/2A & \text{при } -A \leq x \leq A, \\ 0 & \text{при } |x| > A. \end{cases}$$

О т в е т.

$$w_r(r) = \begin{cases} \pi r/2A^2 & \text{при } r \leq A, \\ \frac{\pi r}{2A^2} \left(1 - \frac{4}{\pi} \arccos \frac{A}{r}\right) & \text{при } A < r \leq \sqrt{2}A, \\ 0 & \text{при } \sqrt{2}A < r. \end{cases}$$

13. Коэффициент  $B$  определен в (26.2) через  $\overline{(x-y)^2}$ , т. е. через среднее квадратичное отклонение случайной конечной точки  $x$  в момент  $t$  от фиксированной исходной точки  $y$  в момент  $t-\tau$ . Переход из  $y$  в  $x$  обусловлен как хаотическим движением (диффузией), так и систематическим сносом со скоростью

$$A(y, t) = \lim_{\tau \rightarrow 0} \frac{\overline{x-y}}{\tau} = \lim_{\tau \rightarrow 0} \frac{\bar{x}-y}{\tau},$$

где  $\bar{x}$  — среднее значение  $x$  в момент  $t$ . Казалось бы естественным ввести в качестве меры чисто случайного (диффузионного) смещения не величину  $B$ , а

$$\tilde{B} = \lim_{\tau \rightarrow 0} \frac{\overline{(x-\bar{x})^2}}{\tau}.$$

Показать, что  $\tilde{B} = B$ , и истолковать этот результат.

Р е ш е н и е. С точностью до первого порядка относительно  $\tau$  имеем

$$\bar{x} - y \approx A(y, t)\tau.$$

Поэтому

$$\begin{aligned} \tilde{B}(y, t) &= \lim_{\tau \rightarrow 0} \frac{\overline{(x-\bar{x})^2}}{\tau} = \lim_{\tau \rightarrow 0} \left\{ \frac{1}{\tau} \overline{(x-y-A\tau)^2} \right\} = \\ &= \lim_{\tau \rightarrow 0} \left\{ \frac{1}{\tau} [\overline{(x-y)^2} - 2A\tau \overline{(x-y)} + A^2\tau^2] \right\} = \\ &= \lim_{\tau \rightarrow 0} \left\{ \frac{1}{\tau} [\overline{(x-y)^2} - A^2\tau^2] \right\} = B(y, t). \end{aligned}$$

Причина, очевидно, в том, что полное отклонение  $x-y$  и чисто случайное  $x-\bar{x}$ , оба растущие по диффузионному закону (стандарты обоих порядка  $\sqrt{\tau}$ ), отличаются друг от друга в более высоком (первом) порядке по  $\tau$ :

$$(x-y) - (x-\bar{x}) = \bar{x} - y \approx A\tau.$$

14. Уравнение Колмогорова (23.8) для случайного процесса с дискретными состояниями было получено из уравнения Смолуховского в предположении, что существуют вероятности перехода за единицу времени  $A_{jk}(t)$  ( $k \neq j$ ), причем  $A_{jj}(t) = -\sum_{k \neq j} A_{jk}(t)$ . При этом предположении уравнение (23.8) можно вывести и из уравнения Маркова. Показать это для одношаговых случайных блужданий, у которых возможны смещения на один шаг вправо и влево (с вероятностями  $p$  и  $q$ ), а также «шаг на месте» с вероятностью  $\bar{p}$ .

Решение. Уравнение Маркова для вероятности перехода  $p_{N,m}$  будет

$$p_{N+1,m} = p \cdot p_{N,m-1} + q \cdot p_{N,m+1} + \tilde{p} \cdot p_{N,m}, \quad p + q + \tilde{p} = 1.$$

Полагая  $N \Delta t = t$ , перепишем это уравнение в виде

$$p_m(t + \Delta t) = p \cdot p_{m-1}(t) + q \cdot p_{m+1}(t) + \tilde{p} \cdot p_m(t).$$

Если

$$p = A_{m-1,m} \Delta t + o(\Delta t), \quad q = A_{m+1,m} \Delta t + o(\Delta t), \quad \tilde{p} = 1 + A_{m,m} \Delta t + o(\Delta t)$$

причем

$$A_{m,m} = -(A_{m-1,m} + A_{m+1,m}),$$

то в пределе при  $\Delta t \rightarrow 0$  получаем уравнение Колмогорова:

$$\dot{p}_m(t) = A_{m-1,m} p_{m-1}(t) + A_{m+1,m} p_{m+1}(t) - (A_{m-1,m} + A_{m+1,m}) p_m(t).$$

В частности, если блуждания симметричны ( $A_{m-1,m} = A_{m+1,m}$ ), то, выбирая такой масштаб для  $t$ , чтобы эти временные плотности вероятности были равны единице, получаем

$$\dot{p}_m = p_{m-1} + p_{m+1} - 2p_m.$$

Под  $p_m$  можно, конечно, понимать как вероятность перехода  $p(t, ma|0, 0)$ , так и вероятность состояния  $p(t, ma)$ . Следует обратить внимание на то, что для существования  $A_{j,m}$  введение «шага на месте» необходимо.

15. В § 24 уравнение Эйнштейна — Фоккера для процесса с непрерывными возможными состояниями было выведено из уравнения Маркова для дискретной последовательности. В предыдущей задаче показано, что предельный переход  $\Delta t \rightarrow 0$  позволяет при определенных условиях получить из уравнения Маркова уравнение Колмогорова для процесса с дискретными состояниями. Естественно ожидать, что последующее «сближение» этих состояний (например, неограниченное укорочение длины шага  $a$  при случайных блужданиях) позволит вывести уравнение Эйнштейна — Фоккера из уравнения Колмогорова. Показать это на примере симметричных одношаговых блужданий, т. е. для уравнения Колмогорова

$$\dot{p}_m = p_{m+1} + p_{m-1} - 2p_m.$$

Решение. Полагая  $x = ma$  и вводя плотность вероятности  $v(t, x) = p(t, ma)/a = p_m/a$ , приводим уравнение Колмогорова к виду

$$\dot{v}(t, x) = v(t, x + a) + v(t, x - a) - 2v(t, x).$$

Разложение  $v(t, x \pm a)$  по степеням  $a$  дает

$$\dot{v}(t, x) = a^2 \frac{\partial^2 v}{\partial x^2} + \frac{a^4}{12} \frac{\partial^4 v}{\partial x^4} + \dots$$

Если вместо  $t$  ввести  $\tau = a^2 t$  и разложить  $v(\tau, x, a^2)$  по степеням  $a^2$ :

$$v(\tau, x, a^2) = v^{(0)}(\tau, x) + a^2 v^{(1)}(\tau, x) + \dots,$$

то получим систему уравнений последовательных приближений:

$$\frac{\partial v^{(0)}}{\partial \tau} = \frac{\partial^2 v^{(0)}}{\partial x^2}, \quad \frac{\partial v^{(1)}}{\partial \tau} = \frac{\partial^2 v^{(1)}}{\partial x^2} + \frac{1}{12} \frac{\partial^4 v^{(0)}}{\partial x^4}, \dots$$

Нулевое приближение, т. е. уравнение для  $v^{(0)}$ , представляет собой уравнение Эйнштейна — Фоккера.

Учет «поправок на дискретность»  $v^{(1)}$ ,  $v^{(2)}$ , ... представляет интерес во многих задачах, в которых описание дискретной случайной структуры или

процесса заменяется «сглаженным» описанием при помощи дифференциальных уравнений в частных производных для непрерывных переменных. Иногда важно выяснить, как сказывается на таких уравнениях многошаговость исходной дискретной задачи и к каким граничным условиям для непрерывных переменных приводят (в разных порядках по  $a^2$ ) различные способы ограничения области изменения дискретных переменных. Эти вопросы подробно исследованы в работе [29] на моделях одношаговых и двухшаговых случайных блужданий.

16. Пользуясь характеристической функцией условного распределения  $\tilde{v}(t, x | t - \tau, y; T, X)$ , показать, что из интегрального соотношения (26.16) следует уравнение вида (26.17) для условной вероятности  $v(t, x | T, X)$ , коэффициенты которого  $A_n$  определены по (26.18) [12].

Решение. Запишем характеристическую функцию в виде

$$\begin{aligned}\Phi(t, u | t - \tau, y; T, X) &= \int_{-\infty}^{+\infty} e^{iu(x-y)} \tilde{v}(t, x | t - \tau, y; T, X) dx = \\ &= \sum_{n=0}^{\infty} \frac{(iu)^n}{n!} a_n(t - \tau, y; T, X),\end{aligned}$$

где

$$a_n(t - \tau, y; T, X) = \int_{-\infty}^{+\infty} (x - y)^n \tilde{v}(t, x | t - \tau, y; T, X) dx.$$

Обратное преобразование Фурье дает

$$\begin{aligned}\tilde{v}(t, x | t - \tau, y; T, X) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iu(x-y)} \Phi(t, u | t - \tau, y; T, X) du = \\ &= \sum_{n=0}^{\infty} \frac{a_n}{2\pi n!} \int_{-\infty}^{+\infty} e^{-iu(x-y)} (iu)^n du = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} a_n \frac{\partial^n \delta(x-y)}{\partial x^n}.\end{aligned}$$

Подставив это выражение для  $\tilde{v}$  в интеграл в (26.16), получаем

$$\begin{aligned}v(t, x | T, X) &= \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int_{-\infty}^{+\infty} a_n(y, t - \tau; T, X) v(t - \tau, y | T, X) \frac{\partial^n \delta(x-y)}{\partial x^n} dy = \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial x^n} [a_n(x, t - \tau; T, X) v(t - \tau, x | T, X)]\end{aligned}$$

Член суммы с  $n = 0$  равен просто  $v(t - \tau, x | T, X)$ . Перенесем его в левую часть, разделим все уравнение на  $\tau$  и перейдем к пределу при  $\tau \rightarrow 0$ . Учитывая, что, согласно (26.18),

$$\lim_{\tau \rightarrow 0} \frac{a_n(x, t - \tau; T, X)}{\tau} = A_n(x, t | T, X),$$

получаем

$$\frac{\partial v(t, x | T, X)}{\partial t} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial x^n} [A_n(x, t | T, X) v(t, x | T, X)].$$

Очевидно, весь проделанный вывод остается в силе и тогда, когда  $\tilde{v} = v$ , а условия  $(T, X)$  сводятся к предшествующему состоянию  $x_0 = x(t_0)$  самого процесса  $x(t)$ , т. е. в случае марковского процесса. Мы получаем тогда уравнение (26.4) для вероятности перехода  $v(t, x | t_0, x_0)$ , поскольку отличны от нуля только  $A_1$  и  $A_2$ .

17. Показать, что плотность вероятности  $v(x | t, x_0)$ , выражающаяся формулой (31.5), удовлетворяет уравнению Смолуховского

$$v(x | t, x_0) = \int_{-\infty}^{+\infty} v(x | t - \theta, y) v(y | \theta, x_0) dy, \quad 0 < \theta < t.$$

Решение. Разлагая  $t^n = [(t - \theta) + \theta]^n$  по формуле бинома, нетрудно убедиться, что для закона Пуассона  $P(n) \equiv P_n(t) = \frac{(n_1 t)^n}{n!} e^{-n_1 t}$  справедливо соотношение

$$P_n(t) = \sum_{k=0}^n P_{n-k}(t - \theta) P_k(\theta). \quad (1)$$

Далее, из выражения (31.2) для  $p_n(x | x_0)$  вытекает формула

$$\int_{-\infty}^{+\infty} p_m(x | y) p_k(y | x_0) dy = p_{m+k}(x | x_0), \quad (2)$$

справедливая и в том случае, когда  $m$  и (или)  $k$  — нули.

Записав  $v(x | t, x_0)$  в виде (31.4):

$$v(x | t, x_0) = \sum_{n=0}^{\infty} P_n(t) p_n(x | x_0),$$

получаем

$$v(x | t - \theta, y) v(y | \theta, x_0) = \sum_{k, m=0}^{\infty} P_m(t - \theta) P_k(\theta) p_m(x | y) p_k(y | x_0).$$

Интегрируя это выражение по  $y$  и учитывая (2), находим

$$\begin{aligned} & \int_{-\infty}^{+\infty} v(x | t - \theta, y) v(y | \theta, x_0) dy = \\ &= \sum_{k, m=0}^{\infty} P_m(t - \theta) P_k(\theta) p_{m+k}(x | x_0) = \sum_{n=0}^{\infty} P_n(x | x_0) \sum_{k=0}^n P_{n-k}(t - \theta) P_k(\theta). \end{aligned}$$

В силу (1) это равно  $v(x | t, x_0)$ .

18. Найти вероятность перехода, одномерную плотность вероятности и функцию корреляции в том частном случае пуассоновского процесса (31.5), когда апостериорная (после скачка) плотность вероятности  $\varphi(x_v | x_{v-1}) dx_v$  не зависит от  $x_{v-1}$ , т. е. равна  $\varphi(x_v) dx_v$ . Если  $x_v$  — скорость частицы в газе, то данный случай отвечает легкой частице (например, электрону), соударяющейся с тяжелыми частицами. Легкая частица практически «забывает» свою скорость до удара при каждом соударении.

**Решение.** Согласно (31.2) имеем  $p_n(x|x_0) = \varphi(x)$ , и поэтому вероятность перехода (31.5) равна

$$\begin{aligned} v(x|t, x_0) &= e^{-n_1 t} \left\{ \delta(x - x_0) + \varphi(x) \sum_{n=1}^{\infty} \frac{(n_1 t)^n}{n!} \right\} = \\ &= e^{-n_1 t} \delta(x - x_0) + (1 - e^{-n_1 t}) \varphi(x), \end{aligned} \quad (1)$$

т. е.  $x = x_0$  с вероятностью  $e^{-n_1 t}$  того, что в  $(0, t)$  не было соударений, или же  $x$  имеет распределение  $\varphi(x)$  с вероятностью  $1 - e^{-n_1 t}$  того, что в  $(0, t)$  было хотя бы одно соударение. Такая эволюция распределения (1) описывает, например, поведение скоростей в пучке легких частиц, проходящем через газ тяжелых частиц. Слагаемое с  $\delta(x - x_0)$  описывает частицы, не испытавшие соударений до момента  $t$  и поэтому все еще принадлежащие пучку, а второе слагаемое — частицы, рассеянные газом.

При  $t \rightarrow \infty$  устанавливается стационарное распределение  $w_1(x) = \varphi(x)$ , и, следовательно, двумерное распределение будет

$$\begin{aligned} w_2(x, t; x_0, 0) &= w_1(x_0) v(x|t, x_0) = \\ &= \varphi(x_0) [e^{-n_1 t} \delta(x - x_0) + (1 - e^{-n_1 t}) \varphi(x)]. \end{aligned}$$

Вычисляя при помощи  $w_2$  функцию корреляции, получаем

$$\psi(t) = \overline{xx_0} - \bar{x} \cdot \bar{x}_0 = \int xx_0 w_2(x, t; x_0, 0) dx dx_0 - \bar{x}^2 = \sigma^2 e^{-n_1 t},$$

где  $\sigma^2 = \overline{x^2} - \bar{x}^2$ . Функция корреляции такая же, как у нормального марковского процесса (см. задачу 5), хотя распределение  $w_1$  — даже при гауссовом распределении  $\varphi(x)$  — не является нормальным.

**19.** Найти среднее время достижения границ при наличии постоянной систематической скорости  $A$ , направленной к верхней границе  $b > x$  (т. е.  $A > 0$ ), для случаев, когда имеется нижняя граница  $a < x$  и когда она отсутствует. Коэффициент диффузии постоянен.

**Решение.** Уравнение (32.10) при постоянных  $A$  и  $B$

$$\frac{B}{2} \bar{t}'' + A \bar{t}' + 1 = 0$$

имеет общее решение

$$\bar{t}(x) = \frac{1}{A} \left( C_1 - x + \frac{1}{\lambda} \right) + C_2 e^{-\lambda x}, \quad \lambda = \frac{2A}{B}. \quad (1)$$

При наличии обеих границ, находя  $C_1$  и  $C_2$  из условий  $\bar{t}(a) = \bar{t}(b) = 0$ , получаем

$$\bar{t}(x) = \frac{1}{A} \left[ \frac{ae^{-\lambda b} - be^{-\lambda a} + (b-a)e^{-\lambda x}}{e^{-\lambda b} - e^{-\lambda a}} - x \right].$$

При  $A = 0$  (а значит, и  $\lambda = 0$ ) мы возвращаемся к формуле (31.15).

Если при  $A \neq 0$  отодвинуть нижнюю границу в бесконечность ( $a \rightarrow -\infty$ ), то получим

$$\bar{t}(x) = \frac{b-x}{A},$$

т. е. среднее время достижения верхней границы  $b$  равно в этом случае просто времени перемещения до нее с систематической скоростью  $A$ . Это следует и из общего решения (1), если наложить только одно условие  $\bar{t}(b) = 0$  и потребовать неотрицательности  $\bar{t}(x)$ . Это последнее требование сводится, как нетрудно видеть, к условию наиболее медленного роста  $\bar{t}(x)$  при  $x \rightarrow -\infty$  (т. е. при удалении начальной точки  $x$  от границы  $b$ ), что достигается при отсутствии в (1) экспоненциального члена ( $C_2 = 0$ ).

## СТОХАСТИЧЕСКИЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

### § 33. Постановка вопроса

Мы уже отмечали, что марковские процессы можно рассматривать как непосредственное обобщение детерминированных процессов, а именно как отклик динамической системы на случайное воздействие. Конечно, это не означает, что отклик любой динамической системы на произвольное случайное воздействие обязательно будет процессом без вероятностного последствия. Напротив, можно предвидеть заранее, что для этого необходимо выполнение определенных условий. Однако в рамках того чисто вероятностного подхода, который был изложен в предыдущей главе, т. е. при выводе уравнений для функций распределения (вероятностей состояний, вероятностей перехода), мы, по существу, не анализировали вопроса ни о динамических свойствах системы, ни о вероятностных свойствах воздействий на нее. Применяя теорию марковских процессов (например, к автогенератору), мы были вынуждены поэтому апеллировать к различным качественным соображениям об условиях марковости отклика. Действительно, коэффициенты  $A_{jk}(t)$  в уравнениях Колмогорова для марковских процессов с дискретными состояниями (§ 23), коэффициенты  $A(x, t)$  и  $B(x, t)$  в уравнении Эйнштейна — Фоккера для диффузионных марковских процессов (§§ 26, 27), а также ядро  $\varphi(x|t, x_0)$  и вероятность  $a(t, x)$  в уравнении Колмогорова — Феллера в случае скачкообразного марковского процесса (§ 31) — все эти величины появились в результате определенных предположений о том, как должны вести себя искомые вероятности перехода на бесконечно малых промежутках времени. Связь этих функций с динамическими характеристиками системы и с вероятностными свойствами случайных воздействий установлена не была.

Обратимся теперь к другому подходу к случайным процессам в динамических системах, при котором исходными являются

уравнения не для вероятностей перехода (или других функций распределения), а для самого случайного отклика  $\xi(t)$ . Этот подход был введен Ланжевеном в связи с теорией броуновского движения [1], но его значение и возможности значительно шире. Достаточно указать на то, что он не ограничен требованием марковости отклика  $\xi(t)$ .

Для пояснения постановки вопроса возьмем случай *линейной* динамической системы (что тоже не обязательно), описываемой, например, дифференциальным уравнением

$$\hat{L}\xi(t) = f(t). \quad (33.1)$$

Здесь  $f(t)$  — внешняя сила, а  $L$  — линейный дифференциальный оператор, который в общем случае тоже явно зависит от времени  $t$  — через коэффициенты  $a_k(t)$  при производных:

$$\hat{L} = \sum_{k=0}^n a_k(t) \frac{d^k}{dt^k}$$

(*параметрическое воздействие*). Пусть динамическая задача, поставленная уравнением (33.1) с заданными начальными условиями  $\xi^{(k)}(0)$  ( $k = 0, 1, \dots, n-1$ ), имеет единственное решение  $\xi(t)$ . Предположим теперь, что и сила  $f(t)$ , и коэффициенты  $a_k(t)$  и начальные значения  $\xi^{(k)}(0)$  *случайны*, причем почти для всех реализаций этих случайных функций и величин (т. е. почти наверное) остаются в силе условия единственности решения  $\xi(t)$ . Тогда уравнение (33.1) порождает множество реализаций  $\xi(t)$  и естественно ожидать, что все вероятностные свойства этого (теперь уже случайного) процесса  $\xi(t)$  будут вполне определены статистикой исходных случайных функций и величин. Если это справедливо, то можно рассчитывать и на то, что в частном случае, когда отклик  $\xi(t)$  оказывается диффузионным марковским процессом, статистика исходных функций и величин должна *определять* коэффициенты  $A(x, t)$  и  $B(x, t)$  в уравнении Эйнштейна — Фоккера для вероятности перехода  $v(x, t | x_0, t_0)$  случайного процесса  $\xi(t)$ .

Конечно, здесь возникает ряд чисто математических вопросов. Например, в каких случаях можно получить дифференциальные уравнения для  $\omega_n$  или по крайней мере для любых моментов  $\xi(t)$ ? Когда обеспечена марковость процесса  $\xi(t)$ , которая, как сказано, вовсе не обязательна при описанной постановке вопроса? И т. п. Мы не можем и не будем пытаться анализировать эти вопросы во всей полноте, тем более что не все они уже имеют ответ, а ограничимся лишь некоторыми результатами, представляющими непосредственный интерес для приложений.

Прежде всего нам надо познакомиться с еще одним классом случайных функций, который играет важную роль именно в тех случаях, когда отклик  $\xi(t)$  динамической системы оказывается марковским процессом.

### § 34. Случайные функции с независимыми приращениями

Мы уже неоднократно встречались с такими функциями, хотя соответствующий термин еще не применялся. Примерами, с которыми мы уже имели дело, являются: перемещение частицы при брауновском движении как поступательном (§§ 7, 24), так и вращательном (§ 30); флуктуационный набег фазы в ламповом генераторе (§ 29); заряд, перенесенный потоком электронов на анод лампы в условиях, когда отсутствует явление депрессии дробового шума (§§ 5, 10). Во всех этих примерах речь идет о такой случайной функции  $\xi(t)$ , *приращения* которой на *неперекрывающихся* интервалах времени *независимы*.

Разбив промежуток времени  $(0, t)$  на произвольное число  $n$  последовательных интервалов  $\Delta t_i = t_{i+1} - t_i$  ( $t_0 = 0$ ,  $t_n = t$ ), всегда можно поэтому представить значение  $\xi(t)$  как сумму *независимых* слагаемых — значения  $\xi(0)$  и приращений  $\Delta_i \xi = \xi(t_{i+1}) - \xi(t_i)$  функции  $\xi(t)$  на этих последовательных интервалах:

$$\xi(t) = \xi(0) + \sum_{i=0}^{n-1} \Delta_i \xi. \quad (34.1)$$

Это справедливо для пути, пройденного брауновской частицей за время  $t$ , для числа распавшихся за время  $t$  радиоактивных ядер (при условии, что еще можно пренебречь убылью нераспавшихся ядер), для флуктуационного ухода фазы лампового генератора и т. п. Вследствие независимости слагаемых в (34.1), дисперсия  $\xi(t)$  будет

$$D[\xi(t)] = D[\xi(0)] + \sum_{i=0}^{n-1} D[\Delta_i \xi]. \quad (34.2)$$

Если, в частности, рассматриваемая функция с независимыми приращениями  $\xi(t)$  почти наверное *непрерывна*, то при увеличении  $n$  независимые слагаемые  $\Delta_i \xi$  могут быть сделаны сколь угодно малыми. Переходя к пределу при  $n \rightarrow \infty$ , мы можем на основании центральной предельной теоремы заключить, что функция  $\xi(t) - \xi(0)$ , где  $\xi(t)$  — непрерывная функция с независимыми приращениями, распределена нормально с дисперсией

$$\int_0^t B(t) dt, \text{ причем } B(t) > 0.$$



Допустим теперь, что процесс рассматриваемого типа *однороден по  $t$* . Это означает, что распределение  $\Delta_i \xi$  зависит только от длительности интервала  $\Delta t_i = t_{i+1} - t_i$  (но по-прежнему не зависит от значений, принятых функцией в моменты времени  $\theta \leq t_i$ ). Примем, что  $\xi(0) = 0$ . Тогда  $\xi(t)$  равно сумме своих приращений на интервале  $(0, t)$ :

$$\Delta_i \xi = \xi(t_{i+1}) - \xi(t_i) \equiv \eta(t_{i+1} - t_i) = \eta(\Delta t_i),$$

причем  $\xi(t) = \eta(t - 0) = \eta(t)$ . Формулы (34.1) и (34.2) примут вид

$$\xi(t) = \eta(t) = \sum_{i=0}^{n-1} \eta(\Delta t_i), \quad (34.8)$$

$$D[\eta(t)] = \sum_{i=0}^{n-1} D[\eta(\Delta t_i)]. \quad (34.4)$$

Единственное неотрицательное решение функционального уравнения (34.4) — это линейная однородная функция

$$D[\xi(t)] = Bt \quad (B > 0), \quad (34.5)$$

т. е. диффузионный закон. Таким образом, *однородная непрерывная функция с независимыми приращениями распределена по нормальному закону с дисперсией, пропорциональной  $t$* . Мы имели возможность убедиться в этом в § 24, когда перешли от случайной функции с независимыми приращениями, имеющей дискретные возможные значения (скачки на  $\pm a$  с вероятностями, не зависящими от номера скачка), к непрерывному изменению ( $a \rightarrow 0$ ). Если же с самого начала задача ставилась для непрерывной функции, то к тому же результату приводило предположение об однородности случайных толчков ( $B = \text{const}$ , §§ 27, 29).

Обратим теперь внимание на следующий существенный факт: непрерывная функция с независимыми приращениями *не дифференцируема*<sup>1)</sup>. Действительно, дисперсия приращения  $\Delta \xi = \xi(t + \Delta t) - \xi(t)$  — величина порядка  $\Delta t$ :

$$D[\Delta \xi] \sim \Delta t$$

(для однородной функции имеем просто  $D[\Delta \xi] = B \Delta t$ , где  $B = \text{const}$ ). Следовательно,

$$D\left[\frac{\Delta \xi}{\Delta t}\right] = \frac{1}{(\Delta t)^2} D[\Delta \xi] \sim \frac{1}{\Delta t},$$

т. е. в любой момент  $t$  дисперсия нормально распределенной величины  $\Delta \xi / \Delta t$  неограниченно растет при  $\Delta t \rightarrow 0$ , а это и означает,

<sup>1)</sup> Ни в одном из вероятностных смыслов этого понятия (§ 19).

что производной  $d\xi/dt$  не существует ни в одной точке  $t$  (§ 19). Выходит, таким образом, что если пройденный брауновской частицей путь  $s(t)$ , или поступивший на анод лампы заряд  $q(t)$ , или флуктуационный набег фазы в генераторе  $\varphi(t)$  рассматриваются как непрерывные функции с независимыми приращениями, то не существует ни мгновенной скорости  $u = ds/dt$ , ни мгновенной силы тока  $I = dq/dt$ , ни мгновенной частоты  $\omega = d\varphi/dt$ . Между тем физики и инженеры охотно пользуются случайными функциями  $u(t)$ ,  $I(t)$ ,  $\omega(t)$  и т. п., не впадая при этом в ошибки в получаемых результатах. Целесообразно поэтому несколько задержаться на данном вопросе.

Будем для наглядности рассуждать над одномерным поступательным движением брауновской частицы, считая, что  $\xi(t) \equiv s(t)$  — однородная функция с независимыми приращениями. Для простоты примем, что  $\overline{s(t)} = 0$ .

Переход в (27.1) к пределу при  $n \rightarrow \infty$  (§ 19) приводит к выражению  $s(t)$  в виде интеграла

$$s(t) = \int_0^t ds(t'),$$

где  $ds(t')$  — приращение пути на элементе  $(t', t' + dt')$ . Следовательно,

$$D[s(t)] = \overline{s^2(t)} = \int_0^t \int_0^t \overline{ds(t') ds(t'')}. \quad (34.5)$$

Согласно (34.5) двукратный интеграл в правой части должен быть равен  $Bt$ , что возможно, только если функция корреляции для  $ds(t)$  имеет вид

$$\overline{ds(t') ds(t'')} = B \delta(t' - t'') dt' dt''. \quad (34.6)$$

Конечно, это прямое следствие независимости приращений на неперекрывающихся интервалах.

Если формально оперировать мгновенной скоростью  $u(t) = ds/dt$ , то

$$\overline{ds(t') ds(t'')} = \overline{u(t') u(t'')} dt' dt''.$$

Сопоставление с (34.6) показывает, что  $u(t)$  надо при этом считать так называемым *дельта-коррелированным* случайным процессом:

$$\overline{u(t') u(t'')} = B \delta(t' - t''). \quad (34.7)$$

Соответственно средний квадрат  $u(t)$  при всяком  $t$  бесконечно велик. Нетрудно выяснить, в каком пункте возникает такого рода особенность.

Допущение независимости приращений  $\Delta s$  на неперекрывающихся интервалах  $\Delta t$  справедливо при условии, что за время  $\Delta t$  частица испытывает очень много случайных ударов:  $\Delta t \gg \phi$ , где  $\phi$  — среднее время между ударами. Но определение непрерывности  $s(t)$  требует, чтобы при *неограниченном уменьшении*  $\Delta t$  мы имели  $\Delta s \rightarrow 0$  (хотя бы по вероятности). Оба условия строго совместимы только при  $\phi = 0$ , т. е. при *бесконечно-частых* случайных ударах, чем и исключается существование мгновенной скорости  $u = ds/dt$ .

Если же, как это имеет место в действительности,  $\phi \neq 0$ , то за время  $t < \phi$  может не произойти ни одного случайного толчка, а значит,  $s(t)$  уже не будет функцией с независимыми приращениями. Мгновенная скорость  $u(t)$  тогда существует и имеет время корреляции порядка  $\phi$ .

Пусть, например,

$$\overline{u(t')u(t'')} = \frac{B}{2\phi} e^{-|t'-t''|/\phi} \quad (\overline{u^2} = \frac{B}{2\phi}). \quad (34.8)$$

Тогда, поскольку  $\Delta s = \int_t^{t+\Delta t} u(t') dt'$ , имеем

$$\begin{aligned} D[\Delta s] &= \overline{(\Delta s)^2} = \iint_t^{t+\Delta t} \overline{u(t')u(t'')} dt' dt'' = \\ &= 2 \int_t^{t+\Delta t} e^{-t'/\phi} dt' \int_t^{t'} \frac{B}{2\phi} e^{t''/\phi} dt'' = B[\Delta t - \phi(1 - e^{-\Delta t/\phi})]. \end{aligned} \quad (34.9)$$

При  $\Delta t \ll \phi$  получаем

$$\overline{(\Delta s)^2} \approx \frac{B}{2\phi} (\Delta t)^2 = \overline{u^2} (\Delta t)^2, \quad (34.10)$$

так что для всякого  $\Delta t > 0$  неограниченное уменьшение  $\phi$  влечет за собой  $\overline{(\Delta s)^2} \rightarrow \infty$ . Но для  $\Delta t \gg \phi$  функция  $s(t)$  ведет себя так, как *если бы* ее приращения были независимы, и, в частности, как это следует из (34.9), вступает в силу диффузионный закон

$$\overline{(\Delta s)^2} = B \Delta t.$$

Разумеется, эти заключения не связаны со специальным выбором экспоненциальной функции корреляции (34.8) для  $u(t)$ . Существенно лишь то, является ли эта функция корреляции достаточно острой в масштабе интервала  $\Delta t$  (т. е.  $\phi \ll \Delta t$ ) или же этот интервал слишком мал ( $\Delta t \ll \phi$ ).

При учете отличного от нуля времени корреляции  $\phi$  мы не можем переходить в (34.2) к пределу при  $n \rightarrow \infty$ , считая при

этом, что приращения  $\Delta_i \xi$  на интервалах времени  $\Delta t_i$  остаются независимыми. Увеличение  $n$  ограничено условием, что интервалы  $\Delta t_i \sim t/n$  еще велики по сравнению с  $\vartheta$ , т. е.  $n$  ограничено и сверху ( $1 \ll n \ll t/\vartheta$ ). Тем самым, и нормальность распределения функции  $\xi(t) - \xi(0)$  имеет место лишь приближенно, но тем точнее, чем меньше  $\vartheta$ .

Заметим теперь, что условие  $\Delta t_i \gg \vartheta$  для брауновской частицы означает  $\Delta t \gg 10^{-21}$  сек. Таким образом, для промежутков времени, которые в макроскопических масштабах вполне могут считаться бесконечно малыми (скажем,  $10^{-12}$  сек), число случайных ударов еще огромно ( $\sim 10^9$ ). Всякий раз, когда имеет место такое положение вещей, т. е. когда времена порядка  $\vartheta$  в реальных условиях неуловимы и никак не проявляются в наблюдаемых явлениях, практически все происходит так, как будто  $\vartheta = 0$ <sup>1)</sup>. Наглядная и удобная идеализация  $\vartheta = 0$  позволяет рассматривать  $s(t)$  как непрерывную функцию с независимыми приращениями и вместе с тем, если помнить, что на самом деле существует некоторое достаточно малое  $\vartheta \neq 0$ , позволяет не избегать мгновенной скорости  $u = ds/dt$ , приписывая ей дельта-корреляцию (34.7). В сущности, это не связано ни с какими математическими затруднениями, так как одно из основных «свойств»  $\delta$ -функции — «работать» только под знаком интеграла.

### § 35. Простой пример стохастического дифференциального уравнения

Вернемся к динамическому уравнению первого порядка (система с 1/2 степени свободы), примером которого было уравнение для малых флуктуаций амплитуды в автогенераторе [первая формула (29.1)], т. е. уравнению вида

$$\dot{x} = -\beta x. \quad (35.1)$$

С таким же уравнением мы имеем дело в задачах о скорости  $u$  одномерного движения частицы массы  $m$  в среде с вязким трением  $F = -hu$  ( $\beta = h/m$ ), или о смещении  $s$  этой частицы, но лишенной массы и привязанной к пружине с коэффициентом упругости  $k$  ( $\beta = k/h$ ), или о напряжении  $V$  на емкости  $RC$ -контура ( $\beta = 1/RC$ ), или о токе  $I$  в  $LR$ -контуре ( $\beta = R/L$ ) и т. д.

В соответствии со сказанным в § 28, мы рассчитываем на то, что при действии на динамическую систему (35.1) достаточно «густых» (по сравнению со временем установления  $1/\beta$ ) однородных толчков отклик  $x(t)$  будет непрерывным однородным

<sup>1)</sup> Со спектральной точки зрения это означает, что верхняя граница используемых частот лежит гораздо ниже  $1/\vartheta$  (см. § 43).

марковским процессом с вероятностью перехода  $v(x|t, x_0)$ , удовлетворяющей уравнению Эйнштейна — Фоккера

$$\frac{\partial v}{\partial t} = \beta \frac{\partial x v}{\partial x} + \frac{B}{2} \frac{\partial^2 v}{\partial x^2}, \quad (35.2)$$

т. е. уравнению (29.2), но в одномерном случае, когда нет зависимости  $v$  от второй переменной. По способу, мотивированному в § 28, коэффициент  $A(x, t)$  в (35.2) приравнен выражению для  $\dot{x}$ , т. е. правой части уравнения (35.1):  $A = -\beta x$ .

При начальном условии

$$v(x|0, x_0) = \delta(x - x_0)$$

решение уравнения (35.2) выражается нормальным законом

$$v(x|t, x_0) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\bar{x})^2/2\sigma^2}, \quad (35.3)$$

где

$$\bar{x} = x_0 e^{-\beta t}, \quad \sigma^2 = \frac{B}{2\beta} (1 - e^{-2\beta t}) \quad (35.4)$$

[см. (29.5) и (29.6)]. В пределе при  $t \rightarrow \infty$ , т. е. для  $t \gg 1/\beta$ , формула (35.3) переходит в не зависящее от  $x_0$  стационарное распределение  $w_1(x)$  с  $\bar{x} = 0$  и  $\sigma_\infty^2 = B/2\beta$ . В задаче о скорости  $u$  частицы в вязкой среде, когда  $x \equiv u$  и  $\beta = \hbar/m$ , распределение  $w_1(u)$  должно быть максвелловским:

$$w_1(u) = \sqrt{\frac{m}{2\pi kT}} e^{-mu^2/2kT},$$

так что  $\sigma_\infty^2 = kT/m = B/2\beta$ , откуда  $B = 2kTh/m^2$ . Аналогичные выражения для  $B$  можно написать и в остальных перечисленных выше задачах — просто как следствие теоремы о равнораспределении энергии по степеням свободы: средняя энергия системы с  $1/2$  степени свободы должна быть равна  $kT/2$  (в данном случае  $m\bar{u}^2/2 = m\sigma_\infty^2/2 = kT/2$ ).

Такова при сделанных исходных допущениях чисто вероятностная схема решения задачи о флуктуациях. Теперь мы поступим иначе. Введем в уравнение (35.1) случайную (или флуктуационную) силу  $f(t)$ :

$$\dot{x} = -\beta x + f(t). \quad (35.5)$$

Если для конкретности рассуждать над задачей о движении частицы в неограниченной вязкой среде, то речь идет об уравнении движения

$$m\dot{u} = -\hbar u + \tilde{F}(t), \quad (35.6)$$

в котором воздействие среды на частицу разбито на две части: систематическую силу трения  $F = -\hbar u$  и случайную силу  $F(t)$ .

Предполагая, что систематическая сила трения выражается законом Стокса (для сферической частицы радиуса  $a$  имеем  $h = 6\pi a\eta$ , где  $\eta$  — вязкость жидкости), мы делаем два допущения.

Во-первых, должно быть выполнено условие ламинарности обтекания частицы, т. е. малости числа Рейнольдса:

$$\text{Re} = \frac{\rho a u}{\eta},$$

где  $\rho$  — плотность жидкости. Если для  $u$  взять значение средней квадратичной скорости теплового движения [ $u \sim (kT/\rho_1 a^3)^{1/2}$ ,  $\rho_1$  — плотность вещества частицы], т. е. учесть самые быстрые дрожания частицы, то

$$\text{Re} = \frac{\rho}{\eta} \left( \frac{kT}{\rho_1 a} \right)^{1/2}.$$

При  $\eta/\rho \sim 10^{-2}$ ,  $\rho \sim \rho_1 \sim 1$ ,  $kT \sim 10^{-13}$  имеем  $\text{Re} \sim 3 \cdot 10^{-5} a^{-1/2}$ , что даже для молекулярных размеров  $a \sim 10^{-8}$  дает значение  $\text{Re} \sim 0,3$ . Таким образом, условие ламинарности выполнено.

Во-вторых, полная систематическая сила, действующая на шар, движущийся в вязкой несжимаемой жидкости, равна, согласно Буссине,

$$F = -hu - m'\dot{u} - 6\sqrt{\frac{hm'}{\pi}} \int_0^\infty \dot{u}(t - \alpha^2) d\alpha,$$

где  $m' = 2\pi a^3 \rho/3$  — присоединенная масса, равная половине массы, вытесненной частицей жидкости. В уравнении (35.6) из полной силы  $F$  удержан только первый член. Но при  $\rho \sim \rho_1$  второй и третий члены одного порядка с  $m'\dot{u}$ . В отношении  $m'\dot{u}$  это несущественно, так как роль этого члена сводится лишь к изменению эффективной массы частицы. Более важен третий член, выражающий вязкое гидродинамическое последствие (см. §§ 15 и 21), при учете которого система приобретает бесконечное множество степеней свободы.

При наличии вязкого (а тем самым и вероятностного) последствия средний квадрат смещения частицы был найден В. В. Владимирским и Я. П. Терлецким [2]. Обычное выражение оказывается справедливым лишь для промежутков времени  $t$ , достаточно больших по сравнению со временем релаксации ( $t \gg (m + m')/h$ ). Мы ограничимся упрощенной постановкой задачи, основанной на уравнении (35.5).

Мы будем обращаться с этим стохастическим уравнением так, как если бы это было обычное дифференциальное уравне-

ние. Проинтегрировав его при начальном условии  $x(0) = x_0$ <sup>1)</sup>, получаем

$$x(t) = x_0 e^{-\beta t} + e^{-\beta t} \int_0^t e^{\beta \theta} f(\theta) d\theta. \quad (35.7)$$

Так как по предположению  $\overline{f(t)} = 0$ , усреднение (35.7) по ансамблю случайных сил дает

$$\bar{x} = x_0 e^{-\beta t}, \quad (35.8)$$

т. е. для  $\bar{x}$  получается тот же динамический закон, что и из уравнения (35.1), и из уравнения Эйнштейна — Фоккера (35.2).

Найдем теперь дисперсию  $x(t)$ . Согласно (35.7) и (35.8)

$$D[x(t)] = \overline{(x - \bar{x})^2} = e^{-2\beta t} \int_0^t \int_0^t e^{\beta(\theta + \theta')} \langle f(\theta) f(\theta') \rangle d\theta d\theta', \quad (35.9)$$

и, следовательно, для получения  $D[x]$  надо задать функцию корреляции случайной силы  $\psi_f(\theta, \theta') = \langle f(\theta) f(\theta') \rangle$ . Можно задать *любую* функцию корреляции, допускаемую общими ограничениями ее вида, но мы сделаем специальное предположение, а именно примем, что  $f(t)$  — *стационарный дельта-коррелированный* процесс:

$$\psi_f(\theta - \theta') = \langle f(\theta) f(\theta') \rangle = C \delta(\theta - \theta'), \quad (35.10)$$

где  $C$  — постоянная. Заметим, что тем самым импульс силы

$$p(t) = \int_0^t f(\theta) d\theta$$

представляет собой непрерывную случайную функцию с независимыми приращениями и, следовательно, распределен нормально при любом  $t$  (§ 34).

Подставив (35.10) в (35.9), находим

$$\begin{aligned} \sigma^2 = D[x] &= C e^{-2\beta t} \int_0^t \int_0^t e^{\beta(\theta + \theta')} \delta(\theta - \theta') d\theta d\theta' = \\ &= C e^{-2\beta t} \int_0^t e^{2\beta \theta} d\theta = \frac{C}{2\beta} (1 - e^{-2\beta t}). \end{aligned} \quad (35.11)$$

<sup>1)</sup> Здесь и далее мы будем считать начальные условия детерминированными. В противном случае  $x_0$  было бы случайной величиной, заданной своим распределением  $w(x_0)dx_0$ , и вычисляемые далее условные моменты (т. е. моменты при фиксированном значении  $x_0$ ) надо было бы усреднять еще и по этому распределению.

Если положить  $C = B$ , то это совпадает с выражением (35.4) для  $\sigma^2$ , полученным из уравнения Эйнштейна — Фоккера (35.2).

Мы нашли только моменты  $\bar{x}$  и  $D[x]$ , но можно утверждать больше. Поскольку приращение импульса  $\Delta p = f(t)\Delta t$  распределено при всяком  $\Delta t$  нормально, постольку разность  $x - \bar{x}$  представляет собою, согласно (35.7), сумму (или, точнее, предел суммы) нормально распределенных величин. Следовательно, распределение  $x - \bar{x}$  тоже дается гауссовым законом с дисперсией (35.11). Это *условное* распределение (при условии  $x(0) = x_0$ ), если принять  $C = B$ , просто совпадает с (35.3). Далее, нетрудно убедиться прямой подстановкой, что такого вида условные вероятности удовлетворяют уравнению Смолуховского (являются вероятностями перехода), т. е. процесс  $x(t)$  оказывается марковским. Таким образом, если в стохастическом дифференциальном уравнении (35.5) случайная сила  $f(t)$  стационарна и дельта-коррелирована [см. (35.10)], то отклик  $x(t)$  — диффузионный марковский процесс, у которого вероятность перехода удовлетворяет уравнению Эйнштейна — Фоккера с  $B = C$ .

Оба подхода — основанный на уравнении Эйнштейна — Фоккера и основанный на стохастическом дифференциальном уравнении для случайной функции  $x(t)$  — оказываются в рассмотренной задаче равносильными. Это, конечно, не означает их тождества за пределами этой задачи. Уравнение Эйнштейна — Фоккера обладает, например, несомненным преимуществом в тех случаях, когда наложены определенные ограничения множества возможных значений случайной функции  $x(t)$  (наличие отражающих или поглощающих стенок и т. п.), учитываемые просто соответствующими граничными условиями. При ланжевеновской постановке задачи введение такого рода ограничений довольно сложно. С другой стороны, как это уже было подчеркнуто, ланжевеновский метод не требует, чтобы сила обязательно была дельта-коррелирована.

Стоит, быть может, отметить, что как раз в случае дельта-коррелированной силы оперирование дифференциальным уравнением (35.5) имеет в известном смысле условный характер. Это уравнение написано не для  $\bar{x}$ , а для мгновенного значения  $x(t)$ . Но при бесконечно-частых толчках отклик  $x(t)$  — *не дифференцируемая* функция, т. е.  $\dot{x}(t)$  не существует (ни в каком из вероятностных смыслов понятия производной). Таким образом, все «дифференциальное уравнение» имеет лишь некий *символический* смысл. Это надо понимать следующим образом.

Формальное интегрирование уравнения (35.5) приводит к решению (35.7) для  $x(t)$ , в котором уже нет никаких неприятностей, поскольку оно содержит дельта-коррелированную силу только под интегралом. Другими словами, уравнение (35.5) —



это (в рассматриваемом случае дельта-коррелированной силы) математически некорректная запись для последующего — уже вполне осмысленного и, в конечном счете, единственно интересующего нас — *решения* данного уравнения. Оправданием такого подхода являются хорошо известные преимущества оперирования дифференциальными уравнениями при *постановке* задачи — возможность исходить из общих динамических законов, возможность использования всего существующего арсенала математических средств для получения решения и т. д. Мы не говорим уже о том, что при не дельта-коррелированной  $f(t)$  все оговорки становятся излишними: стохастические дифференциальные уравнения для самих случайных функций приобретают тогда вполне определенное математическое содержание и, сверх того, позволяют выйти за пределы класса марковских процессов.

Постоянная  $C$  в функции корреляции (35.10) характеризует, очевидно, интенсивность случайных толчков. Вернемся к переменным, в которых сила и отклик системы *энергетически сопряжены*, т. е. произведение силы на производную отклика представляет собой мощность, отдаваемую системе. Это справедливо, например, для силы  $F(t)$  в уравнении (35.6), так как отдаваемая частице мощность равна  $F(t)\dot{u}$ . Уравнение (35.6) переходит в (35.5), будучи поделено на массу частицы  $m$ . Таким образом,  $\tilde{f}(t) = \tilde{F}(t)/m$ , так что функция корреляции *настоящей* силы  $\tilde{F}(t)$ , в соответствии с (35.10), равна

$$\langle \tilde{F}(t) \tilde{F}(t') \rangle = m^2 \langle f(t) f(t') \rangle = m^2 C \delta(t - t') = \tilde{C} \delta(t - t').$$

Мы установили выше, что  $C = B$  и что в задаче о скорости брауновской частицы  $B = 2kTh/m^2$ . Следовательно, постоянная  $\tilde{C}$  в функции корреляции силы  $\tilde{F}(t)$  равна

$$\tilde{C} = m^2 C = m^2 B = 2kTh, \quad (35.12)$$

т. е. связана *только с коэффициентом систематического трения*  $h$ . В задаче о токе в  $LR$ -контуре под  $F(t)$  надо понимать случайную тепловую э. д. с.  $e(t)$  (§ 28), а под  $h$  — активное сопротивление контура  $R$ , так что корреляционная постоянная для  $e(t)$  будет  $\tilde{C} = 2kTR$ :

$$\langle e(t) e(t') \rangle = 2kTR \delta(t - t'). \quad (35.13)$$

Метод стохастических дифференциальных уравнений позволяет продвинуться и дальше. Решение (35.7) для  $x$  — это скорость  $u$  в задаче о брауновской частице или сила тока  $I$  в  $LR$ -контуре. Но  $u = \dot{s}$ , а  $I = \dot{q}$ , где  $s$  — пройденный частицей путь, а  $q$  — перенесенный через сечение цепи заряд. Следовательно, мы можем получить статистические характеристики и для этих величин. Будем говорить о  $s$ .

Интегрируя выражение (35.7) для  $u = ds/dt$  при (детерминированном) условии  $s(0) = s_0$ , получаем

$$s - s_0 = \frac{u_0}{\beta} (1 - e^{-\beta t}) + \frac{1}{\beta} \int_0^t [1 - e^{-\beta(t-\theta)}] f(\theta) d\theta. \quad (35.14)$$

Отсюда следует, что

$$\bar{s} = s_0 + \frac{u_0}{\beta} (1 - e^{-\beta t}). \quad (35.15)$$

Очевидно, для величины

$$s - \bar{s} = \frac{1}{\beta} \int_0^t [1 - e^{-\beta(t-\theta)}] f(\theta) d\theta \quad (35.16)$$

справедливы те же соображения, на основании которых мы пришли к выводу, что величина  $x - \bar{x}$ , т. е. в нашем случае  $u - \bar{u}$ , распределена нормально. Таким образом, вероятность перехода для  $s(t)$  есть

$$v(s|t, s_0) = \frac{1}{\sqrt{2\pi D[s]}} \exp \left\{ -\frac{(s - \bar{s})^2}{2D[s]} \right\}.$$

Расчет дисперсии  $s(t)$  проводится так же, как и для  $u(t)$ . В обоих случаях речь идет о случайной функции вида

$$R(t) - \bar{R} = \int_0^t H(t, \theta) f(\theta) d\theta. \quad (35.17)$$

Следовательно, в силу (35.10), имеем

$$\begin{aligned} D[R] &= \overline{(R - \bar{R})^2} = \int_0^t \int_0^t H(t, \theta) H(t, \theta') \overline{f(\theta) f(\theta')} d\theta d\theta' = \\ &= C \int_0^t H^2(t, \theta) d\theta. \end{aligned}$$

Если  $H$  зависит только от разности  $t - \theta$ , то, как нетрудно видеть,

$$D[R] = C \int_0^t H^2(\theta) d\theta. \quad (35.18)$$

Для  $s(t)$  функция  $H$ , согласно (35.16), есть

$$H(\theta) = \frac{1}{\beta} (1 - e^{-\beta\theta}).$$

Подставив это в (35.18) и воспользовавшись (35.12), получаем

$$D[s] = \frac{kT}{\beta^2 m} (2\beta t - 3 + 4e^{-\beta t} - e^{-2\beta t}).$$

При  $\beta t \ll 1$  это дает

$$D[s] = \frac{2kT}{3} \frac{\beta t^3}{m} = \frac{2kTh}{3m^2} t^3.$$

Строго говоря, для брауновской частицы мы не имеем права рассматривать времена  $t \ll 1/\beta$  (времени релаксации), так как задача поставлена без учета вязкого последствия, но в задаче о  $LR$ -контуре это ограничение отпадает. В противоположном случае  $\beta t \gg 1$  имеем

$$D[s] = \frac{2kT}{\beta m} t = \frac{2kT}{h} t \quad (35.19)$$

— формула, полученная впервые Эйнштейном. Она показывает, что оживленность брауновского движения растет с повышением температуры и с уменьшением коэффициента вязкого трения  $h$ . Масса  $m$  в (35.19) не входит, что позволяет получить эту формулу из уравнения для  $s(t)$ , пренебрегая в нем инерцией, т. е. из уравнения

$$h\dot{s} = \tilde{F}(t),$$

откуда

$$s - s_0 = \frac{1}{h} \int_0^t \tilde{F}(t) dt.$$

### § 36. Общий случай уравнения первого порядка и системы таких уравнений при гауссовых дельта-коррелированных воздействиях

Рассмотрим теперь более общий случай *нелинейной* и *неавтономной* системы с  $1/2$  степени свободы (системы первого порядка), описываемой дифференциальным уравнением

$$\dot{x} = \Phi[x, t; F(x, t)], \quad (36.1)$$

где  $\Phi$  — детерминированная функция  $x$ ,  $t$  и воздействия  $F(x, t)$ . Последнее предполагается случайным, вообще говоря, нестационарным процессом, зависящим и от состояния  $x$  самой системы.

Разбив правую часть уравнения (36.1) на ее среднее значение [по ансамблю случайных воздействий  $F(x, t)$ ] и флуктуацию — случайное отклонение от этого среднего:

$$\langle \Phi[x, t; F(x, t)] \rangle = X(x, t), \quad \Phi = \langle \Phi \rangle + f(x, t),$$

всегда можно записать (36.1) в виде уравнения с *аддитивным* случайным воздействием:

$$\dot{x} = X(x, t) + f(x, t), \quad (36.2)$$

где  $X$  — детерминированная функция, а  $\langle f(x, t) \rangle = 0$ .

Вопрос о том, при каких условиях решение  $x(t)$  будет — в данном более общем случае — непрерывным марковским процессом, строго говоря, требует отдельного исследования, но мы ограничимся пока некоторыми наглядными соображениями и фактами, касающимися линейного уравнения (35.5).

Марковость решения уравнения (35.5) приближенно обеспечивается выполнением условия  $\vartheta_x \gg \vartheta_f$ , где  $\vartheta_x = 1/\beta$  — время установления системы, а  $\vartheta_f$  — время корреляции стационарной силы  $f(t)$ . Если считать  $f(t)$  дельта-коррелированной, т. е. положить  $\vartheta = 0$ , то указанное условие выполняется при любом сколь угодно малом  $\vartheta_x$  и процесс  $x(t)$  оказывается марковским вполне строго. Физически весьма правдоподобно, что для марковости решения уравнения (36.2) тоже необходима корреляция  $f(x, t)$ , кратковременная по сравнению с временными масштабами процесса  $x(t)$ . Конечно, понятие времени корреляции  $\vartheta_f$  для *нестационарного* процесса  $f(x, t)$ , равно как и понятие «времени установления»  $\vartheta_x$  для *нелинейной неавтономной* динамической системы, несколько расплывчато. Ясно, однако, что речь идет о *мгновенных* скоростях изменения  $x(t)$  и моментов функции  $f(x, t)$ . Наше условие заключается в том, чтобы при всяком  $t$  можно было выбрать интервал времени  $\tau$  такой, что

$$\vartheta_x \gg \tau \gg \vartheta_f.$$

Таким образом, за макроскопически бесконечно малое время  $\tau$  величина  $x(t)$  практически не меняется, тогда как функция корреляции процесса  $f(x, t)$ , зависящая от  $t$  и  $t' = t + \Delta t$ , обращается в нуль уже при  $\Delta t$  порядка  $\vartheta_f \ll \tau$ . Если  $\vartheta_f = 0$ , т. е. функция корреляции процесса  $f(x, t)$  пропорциональна  $\delta(t - t')$ :

$$\langle f(x, t) f(x', t') \rangle = C(x, x', t) \delta(t - t'), \quad (36.3)$$

то выдвигаемое условие будет выполняться даже при  $\tau \rightarrow 0$  (всюду, где скорость изменения  $x(t)$  конечна, т. е.  $\vartheta_x \neq 0$ ).

Допуская, что этим обеспечивается марковость процесса  $x(t)$ , т. е. что для него существует вероятность перехода  $v(t, x | t_0, x_0)$ , удовлетворяющая уравнению Эйнштейна — Фоккера

$$\frac{\partial v}{\partial t} = - \frac{\partial A(x, t) v}{\partial x} + \frac{1}{2} \frac{\partial^2 B(x, t) v}{\partial x^2}, \quad (36.4)$$

подсчитаем входящие сюда величины  $A(x, t)$  и  $B(x, t)$  в соответствии с их определениями (26.1) и (26.2):

$$A(x, t) = \lim_{\Delta t \rightarrow 0} \frac{\overline{\Delta x}}{\Delta t}, \quad B(x, t) = \lim_{\Delta t \rightarrow 0} \frac{\overline{(\Delta x)^2}}{\Delta t}. \quad (36.5)$$

Усреднение в этих выражениях понималось ранее как усреднение по вероятности перехода  $v(t + \Delta t, x + \Delta x | t, x)$ , но теперь, поскольку мы ожидаем, что стохастическое уравнение (36.2) определяет *все* вероятностные свойства  $x(t)$  [и, в частности, все моменты  $x(t)$ ] через заданные вероятностные характеристики силы  $f(x, t)$ , естественно выполнять усреднение по ансамблю реализаций  $f(x, t)$ .

Согласно (36.2) приращение  $x(t)$  за малое время  $\tau$  есть

$$\Delta x(t, t + \tau) = \int_t^{t+\tau} \{X[x(\theta), \theta] + f[x(\theta), \theta]\} d\theta. \quad (36.6)$$

Как для  $X$ , так и для  $f$  можно написать

$$X[x(\theta), \theta] = X[x(t) + \Delta x(t, \theta), \theta] \approx X[x(t), \theta] + X'_x[x(t), \theta] \cdot \Delta x(t, \theta).$$

Для  $\Delta x(t, \theta)$  уже можно ограничиться первым приближением, т. е. положить в (36.6)  $x(\theta) = x(t)$ :

$$\Delta x(t, \theta) \approx \int_t^\theta \{X[x(t), \theta'] + f[x(t), \theta']\} d\theta'.$$

Подставляя это в предыдущее равенство, а результат — в (36.6), получаем

$$\begin{aligned} \Delta x(t, t + \tau) = & \int_0^{t+\tau} d\theta \{X(x, \theta) + f(x, \theta) + \\ & + \int_t^\theta [X'_x(x, \theta) + f'_x(x, \theta)] [X(x, \theta') + f(x, \theta')] d\theta'\}. \end{aligned}$$

При усреднении этого выражения надо учесть, что при любых  $x = x(t)$  и  $\theta$  мы имеем  $\langle f(x, \theta) \rangle = 0$ , а значит, и  $\langle f'_x(x, \theta) \rangle = 0$ , так что

$$\begin{aligned} \langle \Delta x(t, t + \tau) \rangle = & \int_t^{t+\tau} d\theta \left\{ X(x, \theta) + \right. \\ & \left. + \int_t^\theta [X'_x(x, \theta) X(x, \theta') + \langle f'_x(x, \theta) f(x, \theta') \rangle] d\theta' \right\}. \quad (36.7) \end{aligned}$$

Далее, из (36.3) следует, что

$$\begin{aligned}\langle f'_x(x, \theta) f(x, \theta') \rangle &= \left[ \frac{\partial}{\partial x'} \langle f(x', \theta') f(x, \theta) \rangle \right]_{x'=x} = \\ &= \frac{\partial C(x, x', \theta)}{\partial x'} \Big|_{x'=x} \delta(\theta - \theta').\end{aligned}$$

Так как интеграл от дельта-функции дает на границе интервала значение  $1/2$ , (36.7) принимает вид

$$\begin{aligned}\langle \Delta x(t, t + \tau) \rangle &\approx \int_t^{t+\tau} d\theta \left\{ X(x, \theta) + X'_x(x, \theta) X(x, \theta) (\theta - t) + \right. \\ &\quad \left. + \frac{1}{2} \frac{\partial C(x, x', \theta)}{\partial x'} \Big|_{x'=x} \right\} \approx \left\{ X(x, t) + \frac{1}{2} \frac{\partial C(x, x', t)}{\partial x'} \Big|_{x'=x} \right\} \tau + \\ &\quad + X'_x(x, t) X(x, t) \frac{\tau^2}{2}.\end{aligned}$$

Разделив это на  $\tau$  и переходя к пределу при  $\tau \rightarrow 0$ , в соответствии с (36.5) получаем

$$A(x, t) = X(x, t) + \frac{1}{2} \frac{\partial C(x, x', t)}{\partial x'} \Big|_{x'=x}. \quad (36.8)$$

Заметим, что из (36.3) следует симметрия функции  $C(x, x', t)$  относительно перестановки  $x \rightleftharpoons x'$ . Пользуясь этим, нетрудно показать, что

$$\frac{\partial C(x, x', t)}{\partial x'} \Big|_{x'=x} = \frac{1}{2} \frac{\partial C(x, x, t)}{\partial x},$$

так что  $A(x, t)$  можно записать и иначе:

$$A(x, t) = X(x, t) + \frac{1}{4} \frac{\partial C(x, x, t)}{\partial x}. \quad (36.9)$$

В задачах, в которых случайное воздействие  $f(x, t)$  не зависит от  $x$ , так что множитель при  $\delta(t - t')$  в (36.3) либо функция только от  $t$  (нестационарная сила), либо просто постоянная (стационарная сила), член с производной в (36.8) отсутствует, и мы получаем

$$A(x, t) = X(x, t),$$

т. е.  $A(x, t)$  совпадает с детерминированным членом в правой части стохастического уравнения (36.2), как это и было в рассмотренных ранее примерах.

Имея в виду, что и при вычислении  $B(x, t)$  нас интересуют для  $(\Delta x)^2$  только члены не выше первого порядка по  $\tau$ , мы можем, возводя  $\Delta x$  в квадрат, ограничиться первым приближением в (36.6), положив

$$\Delta x(t, t + \tau) \approx \int_t^{t+\tau} \{X(x, \theta) + f(x, \theta)\} d\theta, \quad x = x(t).$$

Тогда

$$\overline{(\Delta x)^2} = \int_0^{\tau} \int_0^{\tau} \{X(x, \theta) X(x, \theta') + \langle f(x, \theta) f(x, \theta') \rangle\} d\theta d\theta' \approx \\ \approx X^2(x, t) \tau^2 + C(x, x, t) \tau$$

и, следовательно,

$$B(x, t) = C(x, x, t). \quad (36.10)$$

Таким образом, функция  $B(x, t)$  в уравнении (36.4), определяющая диффузионную часть потока вероятности, совпадает с множителем  $C(x, x', t)$  в функции корреляции случайного воздействия (36.3), взятым при  $x' = x$ , т. е. с мерой интенсивности этого воздействия.

Результаты (36.8) и (36.10) получены в предположении, что дельта-корреляция  $f(x, t)$  обеспечивает марковость процесса  $x(t)$ . Но основная идея о том, что стохастические уравнения должны при заданной статистике всех исходных случайных функций и величин полностью определять вероятностные свойства решений, требует большего. Уравнения для функций распределения отклика  $x(t)$  (в общем случае многомерного) и, в частности, марковость  $x(t)$  — если она имеет место — должны *вытекать* из стохастических уравнений. В последнее время в этом направлении получен ряд важных и общих результатов. Для случая системы обыкновенных стохастических дифференциальных уравнений первого порядка с нормальными аддитивными воздействиями, обладающими дельта-корреляцией по времени, математическая теория была развита К. Ито (см. [3] или изложение теории, например в [4]), а для достаточно короткой корреляции — Р. Л. Стратоновичем [5]. Мы будем следовать более простой и физически прозрачной форме теории, принадлежащей В. И. Клячкину и В. И. Татарскому [6, 7], приведя только постановку вопроса и результаты. Доказательство требует более сильных и общих математических средств<sup>1)</sup>.

Пусть  $n$ -мерный случайный процесс  $x(t) = \{x_i(t)\}$ ,  $i = 1, 2, \dots, n$ , удовлетворяет системе стохастических уравнений

$$\dot{x}_i(t) = X_i(x, t) + f_i(x, t), \quad (36.11)$$

где  $X_i$  — детерминированные функции, а  $f_i$  — случайные воздействия, *распределенные нормально* в  $(n+1)$ -мерном пространстве параметров  $(x, t)$  (т. е.  $f_i$  — это случайные поля в этом пространстве), с нулевыми средними значениями ( $\langle f_i(x, t) \rangle = 0$ ).

<sup>1)</sup> Оно опирается на метод, разработанный Е. А. Новиковым [8] для гораздо более общего случая стохастических уравнений в частных производных, т. е. для случайных полей.

и с корреляционной матрицей

$$\psi_{ik}(\mathbf{x}, t; \mathbf{x}', t') = \langle f_i(\mathbf{x}, t) f_k(\mathbf{x}', t') \rangle. \quad (36.12)$$

Пусть отношение характерных времен  $\vartheta_f/\vartheta_x$  [ $\vartheta_f$  — наибольшее из времен корреляции функций  $f_i(\mathbf{x}, t)$ , а  $\vartheta_x$  — наименьшее из времен изменения отклика  $\mathbf{x}(t)$ ] достаточно мало и система такова, что допускает при этом условии отыскание функций распределения  $\mathbf{x}(t)$  методом малых возмущений — в виде разложений по степеням малого параметра  $\vartheta_f/\vartheta_x$ . В нулевом приближении ( $\vartheta_f = 0$ ) корреляционная матрица сил  $f_i(\mathbf{x}, t)$  берется в виде<sup>1)</sup>

$$\psi_{ik}^{\text{эфф}}(\mathbf{x}, t; \mathbf{x}', t') = C_{ik}(\mathbf{x}, \mathbf{x}', t) \delta(t - t'), \quad (36.13)$$

где функции  $C_{ik}$  определены условием равенства интегралов по  $t'$  от истинной и эффективной корреляционных матриц (36.12) и (36.13):

$$C_{ik}(\mathbf{x}, \mathbf{x}', t) = \int_{-\infty}^{+\infty} \psi_{ik}(\mathbf{x}, t; \mathbf{x}', t') dt'. \quad (36.14)$$

Тогда отклик  $\mathbf{x}(t)$  в нулевом приближении (36.13) *оказывается непрерывным (диффузионным) марковским процессом*. Это означает, что многомерные распределения вероятности для  $\mathbf{x}(t)$  выражаются произведениями одномерной плотности вероятности состояния  $w_1(\mathbf{x}, t)$  на цепочку вероятностей перехода  $v(t_k, \mathbf{x}_k | t_{k-1}, \mathbf{x}_{k-1})$ , причем для  $v(t, \mathbf{x} | t_0, \mathbf{x}_0)$  *получается уравнение Эйнштейна — Фоккера*<sup>2)</sup>

$$\frac{\partial v}{\partial t} = - \frac{\partial A_i(\mathbf{x}, t) v}{\partial x_i} + \frac{1}{2} \frac{\partial^2 B_{ik}(\mathbf{x}, t) v}{\partial x_i \partial x_k}, \quad (36.15)$$

в котором

$$A_i(\mathbf{x}, t) = X_i(\mathbf{x}, t) + \frac{1}{2} \frac{\partial C_{ik}(\mathbf{x}, \mathbf{x}', t)}{\partial x'_k} \Big|_{\mathbf{x}'=\mathbf{x}}, \quad (36.16)$$

$$B_{ik}(\mathbf{x}, t) = C_{ik}(\mathbf{x}, \mathbf{x}, t).$$

Для плотности вероятности перехода  $v$  начальное условие есть

$$v(t_0, \mathbf{x} | t_0, \mathbf{x}_0) = \delta(\mathbf{x} - \mathbf{x}_0),$$

<sup>1)</sup> Строго говоря, дельта-корреляция (36.13) несовместима с нормальным распределением, предполагающим конечность дисперсий. Однако в результирующее уравнение Эйнштейна — Фоккера (36.15) входят только интегралы (36.14) от функций корреляции. Поэтому в функциях распределения случайных сил  $f_i$  можно считать дисперсии конечными ( $\vartheta_f \neq 0$ ), но из-за малости  $\vartheta_f$  отклик  $\mathbf{x}(t)$  приближенно такой, как если бы было  $\vartheta_f = 0$ .

<sup>2)</sup> По дважды встречающимся индексам производится суммирование от 1 до  $n$ .



а для плотности вероятности состояния  $w_1$ , удовлетворяющей тому же уравнению (36.15), начальное условие имеет вид

$$w_1(t_0, \mathbf{x}) = w_0(\mathbf{x}),$$

где  $w_0$  — заданное начальное распределение.

Эта общая теорема приводит, таким образом, к уравнению (36.15), исходя непосредственно из стохастических уравнений (36.11). Кроме того, развитый в [6] метод обладает еще и тем преимуществом, что позволяет получать поправки к нулевому (диффузионному) приближению. Поправки возникают двоякого рода. Во-первых, в правой части уравнения (36.15) появляется добавочный член вида  $-\partial S'_k(\mathbf{x}, t)/\partial x_k$ . Это поправка к плотности потока вероятности (порядка не ниже  $\theta_f$ ). Во-вторых, матрица  $C_{ik}$  в (36.16) заменяется на

$$\tilde{C}_{ik}(\mathbf{x}, \mathbf{x}', t) = \int_0^t \psi_{ik}(\mathbf{x}, t; \mathbf{x}', t') dt'.$$

Отличие  $\tilde{C}_{ik}$  от  $C_{ik}$  заметно лишь при  $t \leq \theta_f$ , пренебрежение же членом с  $S'_k$  требует, вообще говоря, ограничения интенсивности флуктуационных сил  $f_i$ . В общем виде выражения для  $S'_k$  довольно громоздки [7], но их исследование в конкретных задачах позволяет уточнять условия применимости диффузионного приближения<sup>1)</sup>.

Мы можем теперь разъяснить высказанное в § 30 утверждение о том, что распространение лучей в случайно-неоднородной среде является марковским процессом только в малоугловом приближении.

Если в качестве независимой переменной берется длина луча  $l$ , то уравнения геометрической оптики записываются в виде

$$\frac{d\mathbf{r}}{dl} = \boldsymbol{\tau}(l), \quad \frac{d\boldsymbol{\tau}}{dl} = \nabla\mu(\mathbf{r}) - \boldsymbol{\tau}(\boldsymbol{\tau}, \nabla\mu) \equiv \mathbf{a}(\mathbf{r}, \boldsymbol{\tau}),$$

где  $\mathbf{r}$  — радиус-вектор точки луча,  $\boldsymbol{\tau}$  — единичный вектор касательной к лучу в этой точке,  $\mu(\mathbf{r}) = \ln n(\mathbf{r})$ ,  $n(\mathbf{r})$  — показатель преломления среды. Если ввести шести-векторы  $\mathbf{f} = \{0, \mathbf{a}\}$ ,  $\mathbf{x} = \{\mathbf{r}, \boldsymbol{\tau}\}$ ,  $\mathbf{X} = \{\boldsymbol{\tau}, 0\}$ , то написанные уравнения приводятся к виду (36.11) с независимой переменной  $l$  (вместо  $t$ ). Пусть  $\mu(\mathbf{r})$  — гауссово случайное поле со средним значением  $\bar{\mu} = 0$ . Казалось бы, введя для  $\mathbf{f}$  функцию корреляции вида (36.13), можно было бы написать для шести-вектора  $\mathbf{x}(l)$ , объединяющего случайные

<sup>1)</sup> Дельта-корреляция (36.13) является лишь необходимым условием марковости процесса  $\mathbf{x}(t)$ .

координаты точек луча  $\mathbf{r}(l)$  и случайные направляющие косинусы луча  $\boldsymbol{\tau}(l)$ , уравнение Эйнштейна — Фоккера (36.15). Но величины  $\mathbf{a}$  зависят от  $\mathbf{x}$  и *не зависят* от  $l$ , так что корреляционная матрица «сил»  $\mathbf{f}$  имеет вид  $\psi_{ik}(\mathbf{x}, \mathbf{x}')$ , т. е. не зависит от  $l$  и  $l'$ . Иначе можно сказать, что интервал корреляции  $\mathbf{f}$  по  $l$  бесконечен, и поэтому записать  $\psi_{ik}$  в виде (36.13) с дельта-функцией  $\delta(l - l')$  невозможно.

Можно, однако, перейти к другому описанию луча, взяв за независимую переменную не его длину  $l$ , а толщину  $z$  пройденного им слоя неоднородной среды (луч падает на слой нормально). Точка луча задается тогда его *поперечным смещением*  $\boldsymbol{\rho} = \{x, y\}$  от точки входа в слой  $x = y = 0$ , а направление — *поперечным вектором*  $\boldsymbol{\tau}_\perp = \{\tau_x, \tau_y\}$ , которые рассматриваются как функции от  $z$ . Уравнения геометрической оптики принимают при этом вид

$$\frac{d\boldsymbol{\rho}(z)}{dz} = \frac{\boldsymbol{\tau}_\perp(z)}{\sqrt{1 - \tau_\perp^2}}, \quad \frac{d\boldsymbol{\tau}_\perp(z)}{dz} = \frac{\mathbf{a}_\perp(\boldsymbol{\rho}, \boldsymbol{\tau}_\perp, z)}{\sqrt{1 - \tau_\perp^2}}, \quad (36.17)$$

где  $\mathbf{a}_\perp = \{a_x, a_y\}$ . Уже отсюда видно, что эти уравнения можно использовать только до первой точки поворота луча на  $90^\circ$  от оси  $z$  (т. е. от его первоначального направления), так как в этой точке  $\sqrt{1 - \tau_\perp^2}$  обращается в нуль. Это значит, что в случайно-неоднородной среде мы должны ограничиться областью, в которой вероятности отрицательных  $\tau_z$  весьма малы, т. е. малыми угловыми отклонениями луча ( $\tau_z \approx 1$ ). В этом *малоугловом приближении* уравнения (36.17) принимают вид

$$\frac{d\boldsymbol{\rho}(z)}{dz} = \boldsymbol{\tau}_\perp(z), \quad \frac{d\boldsymbol{\tau}_\perp(z)}{dz} = \nabla_\perp \mu(\boldsymbol{\rho}, z) \equiv \mathbf{a}_\perp(\boldsymbol{\rho}, z).$$

Четырехмерный случайный процесс  $\mathbf{x}(z) = \{\boldsymbol{\rho}(z), \boldsymbol{\tau}_\perp(z)\}$  возникает под действием *поперечных* случайных «сил»  $\mathbf{f} = \{0, \mathbf{a}_\perp\} = \mathbf{f}(\boldsymbol{\rho}, z)$ , которые зависят теперь от переменной  $z$  и обладают корреляционной матрицей  $\psi_{ik}(\boldsymbol{\rho}, z; \boldsymbol{\rho}', z')$ . Процесс  $\mathbf{x}(z)$  будет марковским в предположении дельта-корреляции неоднородностей среды по  $z$ :

$$\psi_{ik}(\boldsymbol{\rho}, z; \boldsymbol{\rho}', z') = C_{ik}(\boldsymbol{\rho}, \boldsymbol{\rho}', z) \delta(z - z'),$$

что и приводит к уравнению Эйнштейна — Фоккера (27.15), где дополнительно принято, что  $C_{ik} = B\delta_{ik} = \text{const}$ .

Для вращательного брауновского движения сферической частицы подобных ограничений не возникает, так как действующие на частицу случайные вращающие моменты — функции независимой переменной (времени  $t$ ), а от состояния системы (ориентации оси частицы) они вообще не зависят.

### § 37. Стохастические уравнения при случайных воздействиях с произвольными законами распределения

В работе [9] (см. также [10]) получены результаты, обобщающие приведенную в предыдущем параграфе теорему на негауссовы случайные воздействия<sup>1)</sup>. Кроме того, сами стохастические уравнения берутся более общего вида, чем (36.11), а именно в виде системы интегро-дифференциальных уравнений

$$\dot{x}_i = X_i(\mathbf{x}, t) + \int D_{ik}(\mathbf{x}, \mathbf{y}, t) f_k(\mathbf{y}, t) d\mathbf{y}, \quad (37.1)$$

где  $X_i$  и ядро  $D_{ik}$  — детерминированные векторная и тензорная функции. Очевидно, уравнения (37.1) переходят в (36.11), если  $D_{ik}(\mathbf{x}, \mathbf{y}, t) = \delta_{ik} \delta(\mathbf{y} - \mathbf{x})$ . Существенно, что случайные воздействия  $f_k(\mathbf{y}, t)$  входят, как и ранее, *линейно* и что для решения  $\mathbf{x}(t)$  выполняется *принцип причинности*:  $\mathbf{x}(t)$  функционально зависит от  $f_k(\mathbf{x}, \tau)$  только при тех  $\tau$ , которые *предшествуют* моменту времени  $t$  ( $t_0 \leq \tau \leq t$ ). В момент  $t = t_0$  заданы начальные условия  $\mathbf{x}(0) = \mathbf{x}_0$ .

При этих условиях показано, что *точным* следствием уравнений (37.1) является некоторое уравнение для одномерной плотности вероятности  $w_1(\mathbf{x}, t)$ , которое, однако, *не является замкнутым* относительно  $w_1$ , так как содержит член, функционально зависящий от  $f_k(\mathbf{x}, \tau)$  при всех  $t_0 \leq \tau \leq t$ . Только в том случае, когда *воздействия  $f_k(\mathbf{x}, t)$  дельта-коррелированы по  $t$* , указанное уравнение для  $w_1(\mathbf{x}, t)$ <sup>2)</sup> становится *замкнутым* (содержит только одномерную функцию распределения), а процесс  $\mathbf{x}(t)$  оказывается *марковским*:  $n$ -мерная плотность вероятности  $w_n(\mathbf{x}_1, t_1; \dots; \mathbf{x}_n, t_n)$  выражается через  $w_1$  и *вероятность перехода  $v$*  формулой (15.3).

Если дельта-коррелированные  $f_k(\mathbf{x}, t)$  представляют собой гауссов процесс, то, в согласии с приведенной в предыдущем параграфе теоремой, уравнение для  $w_1$  или  $v$  будет уравнением Эйнштейна — Фоккера (*диффузионный марковский процесс*). Если же случайные воздействия — импульсные процессы, состоящие из дельта-импульсов

$$f_k(\mathbf{x}, t) = \sum_{\nu} a_{k\nu} \delta(t - t_{\nu})$$

<sup>1)</sup> В [9] рассмотрены не только случайные процессы, но и случайные поля, но в I части книги мы по-прежнему ограничиваемся процессами.

<sup>2)</sup> Или для условной плотности вероятности  $v(t, \mathbf{x} | t_0, \mathbf{x}_0)$ , если начальное условие при  $t = t_0$  есть

$$v(t_0, \mathbf{x} | t_0, \mathbf{x}_0) = \delta(\mathbf{x} - \mathbf{x}_0).$$

(подчеркнем, что такой процесс дельта-коррелирован, но он не обязательно должен быть пуассоновским или стационарным), то для  $\omega_1$  или  $v$  получается уравнение Колмогорова — Феллера (31.10) (*скачкообразный* марковский процесс).

Обратимся теперь к тому преимуществу стохастических дифференциальных уравнений, о котором уже говорилось выше, а именно возможности рассматривать с их помощью и такие (хотя бы линейные) задачи, когда случайная сила  $f(t)$  не обладает дельта-корреляцией, а значит, отклик системы  $\xi(t)$  уже не является марковским процессом.

Из приведенных выше результатов ясно, что при отказе от дельта-корреляции случайных воздействий общая теория очень усложняется. Мы ограничимся примерами, которые показывают тем не менее, насколько существенные результаты могут быть получены хотя бы при определенных ограничениях. Важно, однако, подчеркнуть, что в число этих ограничений уже не будет включаться в качестве обязательного требования точная или приближенная дельта-корреляция случайных воздействий на динамическую систему.

1. Рассмотрим линейное уравнение первого порядка, описывающее динамическую систему с *параметрическим* воздействием:

$$\dot{x} = [Q(t) + q(t)]x, \quad (37.2)$$

где  $Q(t)$  — детерминированная функция, а  $q(t)$  — случайный процесс с  $\overline{q(t)} = 0$ . Таким образом, это уравнение вида (36.2) с  $X(x, t) = Q(t)x$  и  $f(x, t) = q(t)x$ . Если же ввести функцию  $y = \ln x$ , то для нее мы получим линейное уравнение  $\dot{y} = Q(t) + q(t)$  с аддитивной случайной силой, не зависящей от состояния системы  $y$ . Предполагается, что  $q(t)$  — *нормальный* процесс, но с *произвольной* функцией корреляции

$$\Psi_q(t, t') = \langle q(t) q(t') \rangle. \quad (37.3)$$

Оказывается, что и в этом случае можно без каких-либо дополнительных требований получить из (37.2) замкнутое уравнение для плотности условной вероятности  $v(x, t | x_0, 0)$  решения  $x(t)$ <sup>1)</sup>.

Решение уравнения (37.2) при начальном условии  $x(0) = x_0$  есть

$$x(t) = x_0 \exp \left\{ \int_0^t [Q(\tau) + q(\tau)] d\tau \right\}.$$

<sup>1)</sup> Вывод принадлежит В. И. Татарскому.

Поэтому (условная) характеристическая функция процесса  $x(t)$  запишется в виде

$$\begin{aligned}\varphi(u, t) &= \langle \exp\{iux\} \rangle = \sum_{n=0}^{\infty} \frac{(iu)^n}{n!} \langle x^n(t) \rangle = \\ &= \sum_{n=0}^{\infty} \frac{(iux_0)^n}{n!} \exp\left\{n \int_0^t Q(\tau) d\tau\right\} \left\langle \exp\left\{n \int_0^t q(\tau) d\tau\right\} \right\rangle.\end{aligned}$$

Так как интеграл  $\alpha = n \int_0^t q(\tau) d\tau$  — тоже гауссова величина, а для любой гауссовой величины  $\alpha$  при  $\bar{\alpha} = 0$  справедливо равенство  $\langle e^\alpha \rangle = \exp\{\bar{\alpha}^2/2\}$ , получаем

$$\left\langle \exp\left\{n \int_0^t q(\tau) d\tau\right\} \right\rangle = e^{n^2 F(t)},$$

где

$$F(t) = \frac{1}{2} \iint_0^t \psi_q(\tau_1, \tau_2) d\tau_1 d\tau_2.$$

В результате

$$\varphi(u, t) = \sum_{n=0}^{\infty} \frac{(iux_0)^n}{n!} \exp\left\{n \int_0^t Q(\tau) d\tau + n^2 F(t)\right\}, \quad (37.4)$$

откуда

$$\begin{aligned}\frac{\partial \varphi(u, t)}{\partial t} &= \\ &= \sum_{n=0}^{\infty} \frac{(iux_0)^n}{n!} [nQ(t) + n^2 \dot{F}(t)] \exp\left\{n \int_0^t Q(\tau) d\tau + n^2 F(t)\right\}.\end{aligned} \quad (37.5)$$

Пользуясь тождествами

$$nu^n = \left(u \frac{\partial}{\partial u}\right) u^n, \quad n^2 u^n = \left(u \frac{\partial}{\partial u}\right)^2 u^n$$

и выражением (37.4) для  $\varphi$ , нетрудно преобразовать (37.5) к виду

$$\frac{\partial \varphi(u, t)}{\partial t} = Q(t) \left(u \frac{\partial}{\partial u}\right) \varphi + \dot{F}(t) \left(u \frac{\partial}{\partial u}\right)^2 \varphi. \quad (37.6)$$

Если учесть симметрию корреляционной функции  $\psi_q(\tau_1, \tau_2)$  по отношению к перестановке  $\tau_1 \rightleftharpoons \tau_2$ , то коэффициент  $\dot{F}(t)$  легко

выразить в виде *однократного* интеграла от  $\psi_q$ :

$$\begin{aligned} \dot{F}(t) &= \frac{1}{2} \frac{\partial}{\partial t} \int_0^t d\tau_1 \int_0^t d\tau_2 \psi_q(\tau_1, \tau_2) = \\ &= \frac{1}{2} \int_0^t d\tau_2 \psi_q(t, \tau_2) + \frac{1}{2} \int_0^t d\tau_1 \psi_q(\tau_1, t) = \int_0^t \psi_q(t, \tau) d\tau. \end{aligned} \quad (37.7)$$

Задача, в сущности, уже решена, так как мы получили дифференциальное уравнение для характеристической функции. Для того, чтобы вывести уравнение непосредственно для плотности вероятности

$$v(x, t | x_0, 0) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \varphi(u, t) e^{-iux} du,$$

нужно только выполнить над (37.6) преобразование Фурье. Умножая это уравнение на  $e^{-iux}/2\pi$ , интегрируя по  $u$  и предполагая, что при  $|u| = \infty$  проинтегрированные члены обращаются в нуль, получаем с учетом (37.7)

$$\frac{\partial v}{\partial t} = -Q(t) \frac{\partial x v}{\partial x} + \left[ \int_0^t \psi_q(t, \tau) d\tau \right] \frac{\partial}{\partial x} \left( x \frac{\partial x v}{\partial x} \right). \quad (37.8)$$

Подчеркнем еще раз: ничего, кроме нормальности процесса  $q(t)$ , при выводе уравнения (37.8) не предполагалось, и это уравнение представляет собой *обобщенное* уравнение Эйнштейна — Фоккера вида (26.19), справедливое и для немарковского процесса. Если же случайное воздействие  $q(t)$  дельта-коррелировано:

$$\psi_q(t, \tau) = C(t) \delta(t - \tau), \quad (37.9)$$

то  $\int_0^t \psi_q(t, \tau) d\tau = \frac{1}{2} C(t)$  [половина входит потому, что дельта-выброс лежит на границе интервала  $(0, t]$ ], и мы приходим, согласно изложенной в § 36 общей теореме, к *марковскому* процессу  $x(t)$  с уравнением Эйнштейна — Фоккера

$$\frac{\partial v}{\partial t} = -Q(t) \frac{\partial x v}{\partial x} + \frac{C(t)}{2} \frac{\partial}{\partial x} \left( x \frac{\partial x v}{\partial x} \right). \quad (37.10)$$

Хотя вид уравнений (37.8) и (37.10) одинаков, их смысл различен. Только при *дельта-корреляции* (37.9)  $n$ -мерные функции распределения  $x(t)$  выражаются через произведение одинаковых условных вероятностей  $v$ , т. е. последние являются

вероятностями *перехода* (удовлетворяют уравнению Смолуховского).

Если параметрическое воздействие  $q(t)$  *стационарно*, так что  $\psi_q(t, \tau) = \psi_q(t - \tau)$ , то

$$\int_0^t \psi_q(t - \tau) d\tau = \int_0^t \psi_q(\theta) d\theta$$

и при временах  $t \gg \vartheta_q$  [ $\vartheta_q$  — наибольшее время корреляции  $q(t)$ ], можно устремить верхний предел интеграла в бесконечность. Таким образом, коэффициент в квадратных скобках в (37.8) будет с ростом  $t$  стремиться к постоянному значению, как если бы функция корреляции была  $C\delta(t - \tau)$ . В этом случае уравнение (37.8) примет вид (37.10) с постоянным  $C$  и действительно станет уравнением Эйнштейна — Фоккера. Другими словами, в случае стационарного воздействия  $q(t)$  отклик  $x(t)$  становится с ростом  $t$  марковским процессом. Подчеркнем еще раз, что в действительности  $\vartheta_q \neq 0$  и поэтому процесс будет марковским лишь приближенно — если рассматривать его поведение на интервалах времени  $\Delta t \gg \vartheta_q$ , но не на  $\Delta t \leq \vartheta_q$ .

В противоположность приведенному простому примеру в более общих случаях<sup>1)</sup> получение замкнутого уравнения хотя бы для одномерной вероятности, как уже было сказано, неосуществимо без предположения о дельта-корреляции случайного воздействия.

При наличии же *дельта-корреляции* и при *линейных* исходных уравнениях (линейных относительно  $x_i$ ) всегда можно получить *замкнутые* системы уравнений для моментов любого порядка, т. е. отдельную систему для первых моментов  $\langle x_i(t) \rangle$ , отдельную — для вторых моментов  $\langle x_i(t)x_k(t) \rangle$  и т. д. При указанных условиях эти системы уравнений взаимно не зацепляются. Это было показано уже давно [11] (см. задачу 7) и вновь установлено более общим способом в [10].

2. Возможности продвижения в теории стохастических дифференциальных уравнений зависят, конечно, и от степени сложности рассматриваемой динамической системы. Мы обратимся теперь ко второму примеру, имеющему весьма широкое значение, и именно к *линейным* системам, на которые действуют только *случайные внешние силы*, а параметры системы являются детерминированными.

Речь идет о частном случае уравнений (36.11):

$$\dot{x}_i(t) = a_{ik}(t)x_k(t) + f_i(t), \quad i, k = 1, 2, \dots, n, \quad (37.11)$$

<sup>1)</sup> Например, для системы линейных параметрических уравнений, содержащих параметрические случайные воздействия  $q_{ik}(t)$ :

$$\dot{x}_i = [Q_{ik}(t) + q_{ik}(t)]x_k, \quad i = 1, 2, \dots, n.$$

где  $a_{ik}(t)$  — детерминированные функции, а  $f_i(t)$  — случайные силы, не зависящие от состояния системы  $x = \{x_i\}$ . Никаких специальных предположений о вероятностных свойствах сил  $f_i(t)$  не делается, но они считаются заданными.

Исключая из (37.11) все  $x_i(t)$ , кроме какой-либо одной функции [обозначим ее через  $x(t)$ ], можно сформулировать задачу и в виде одного уравнения  $n$ -го порядка:

$$x^{(n)} + a_1(t) x^{(n-1)} + \dots + a_n(t) = f(t). \quad (37.12)$$

Как уже было подчеркнуто ранее, на первом этапе мы обращаемся со стохастическими уравнениями так же, как если бы это были уравнения для детерминированных функций, т. е. решаем «динамическую» задачу. Именно такой путь полностью открыт в случае линейных уравнений, так как их решение всегда может быть записано в форме некоторого линейного же оператора, действующего на заданную силу  $f(t)$ . Воспользуемся истокообразным представлением решения — интегралом Дюамеля:

$$x(t) = \int_0^t H(t, \theta) f(\theta) d\theta, \quad (37.13)$$

где  $H(t, \theta)$  — так называемый *импульсный отклик* рассматриваемой динамической системы, т. е. ее отклик в момент времени  $t \geq \theta$  на силу  $f(t) = \delta(t - \theta)$  при условии, что до момента  $\theta$  система находилась в покое. Соответственно частное решение (37.13) при произвольной силе  $f(t)$  отвечает нулевым начальным условиям:

$$x^{(k)}(0) = 0, \quad k = 0, 1, \dots, n-1.$$

Если коэффициенты уравнения (37.12) постоянны, т. е. динамическая система является *гармонической*, то  $H(t, \theta) = H(t - \theta)$ .

Из линейности связи (37.13) между  $x(t)$  и  $f(t)$  следует, что моменты  $x(t)$  любого целого порядка выражаются в квадратурах через смешанные моменты того же порядка силы  $f(t)$ . Поскольку последние по предположению известны, мы знаем и *все моменты*  $x(t)$ , что в большинстве практически интересных случаев означает принципиальную возможность построения и конечномерных функций распределения  $x(t)$ . В частности,

$$\langle x(t) \rangle = \int_0^t H(t, \theta) \langle f(\theta) \rangle d\theta, \quad (37.14)$$

$$B(t, t') = \langle x(t) x(t') \rangle = \int_0^t H(t, \theta) d\theta \int_0^{t'} H(t', \theta') B_f(\theta, \theta') d\theta', \quad (37.15)$$



где

$$B_f(t, t') = \langle f(t) f(t') \rangle.$$

В случае гармонической системы интеграл Дюамеля можно записать в виде

$$x(t) = \int_0^t H(t-\theta) f(\theta) d\theta = \int_0^t H(\theta) f(t-\theta) d\theta, \quad (37.16)$$

так что

$$\langle x(t) \rangle = \int_0^t H(\theta) \langle f(t-\theta) \rangle d\theta, \quad (37.17)$$

$$\langle x(t) x(t+\tau) \rangle = \int_0^t H(\theta) d\theta \int_0^{t+\tau} H(\theta') \langle f(t-\theta) f(t+\tau-\theta') \rangle d\theta'. \quad (37.18)$$

Если  $f(t)$  — нормальная случайная функция, то, согласно (37.13), отклик  $x(t)$  тоже будет распределен нормально, поскольку он представляет собой в этом случае сумму (предел суммы) нормально распределенных слагаемых. Так как  $n$ -мерное нормальное распределение содержит в качестве параметров моменты только первого и второго порядков, в этом случае достаточно знать только  $\langle f(t) \rangle$  и  $B_f(t, t')$ . Но в общем случае для нахождения одномерного распределения  $x(t)$  необходимо исчерпывающее задание случайной функции  $f(t)$ , т. е. задание всех ее конечномерных распределений. Мы ограничимся рассмотрением моментов (37.17) и (37.18) в предположении, что  $f(t)$  — *стационарный* случайный процесс с нулевым средним значением:

$$\langle f(t) \rangle = 0, \quad \langle f(t) f(t') \rangle = \psi_f(t-t').$$

Тогда из (37.17) имеем  $\langle x(t) \rangle = 0$ , а из (37.18)

$$\psi(t, \tau) = \int_0^t H(\theta) d\theta \int_0^{t+\tau} H(\theta') \psi_f(\tau + \theta - \theta') d\theta'. \quad (37.19)$$

Таким образом, подынтегральное выражение не зависит от  $t$ , а так как реальные системы обладают «конечной памятью» (с ростом  $\theta$  отклик  $H(\theta)$  стремится к нулю), то из (37.19) ясно, что с увеличением  $t$  функция корреляции  $\psi$  перестает зависеть от  $t$ , т. е. в системе устанавливается стационарный процесс<sup>1)</sup>

<sup>1)</sup> Стационарный в широком смысле, см. § 18.

с функцией корреляции

$$\psi(\tau) = \int_0^{\infty} H(\theta) d\theta \int_0^{\infty} H(\theta') \psi_f(\tau + \theta - \theta') d\theta'.$$

Поскольку импульсный отклик  $H(\theta)$  до момента действия дельта-импульса равен нулю, т. е.  $H(\theta) = 0$  при всех  $\theta < 0$ , можно заменить нижние пределы интегралов на  $-\infty$ . Вводя вместо  $\theta'$  переменную интегрирования  $\chi = \theta' - \theta$ , получаем в результате, что в установившемся режиме

$$\psi(\tau) = \int_{-\infty}^{+\infty} H(\theta) d\theta \int_{-\infty}^{+\infty} H(\theta + \chi) \psi_f(\tau - \chi) d\chi,$$

что можно записать в виде

$$\psi(\tau) = \int_{-\infty}^{+\infty} H_2(\chi) \psi_f(\tau - \chi) d\chi, \quad (37.20)$$

где введена функция

$$H_2(\chi) = \int_{-\infty}^{+\infty} H(\theta) H(\theta + \chi) d\theta. \quad (37.21)$$

Функция  $H_2$ , определяемая импульсным откликом  $H$ , характеризует рассматриваемую динамическую систему, тогда как  $\psi_f$  является характеристикой внешней силы  $f(t)$ .

Заметим, что формула (37.20) — это композиция функций  $H_2$  и  $\psi_f$  в соответствии с определением композиции, данным в § 9. Формула же (37.21) не является композицией, так как переменная интегрирования  $\theta$  входит в оба сомножителя под интегралом с одним и тем же знаком.

Приведем в заключение простую иллюстрацию того, как немарковский процесс в линейной системе может быть приведен к марковскому с более высоким числом измерений, о чем уже упоминалось в § 21.

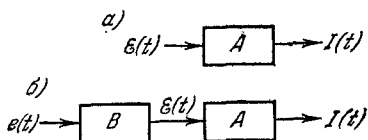


Рис. 27.

Если на вход линейного четырехполюсника  $A$  (рис. 27, а) действует дельта-коррелированная э. д. с.  $\mathcal{E}(t)$ , то выходной ток  $I(t)$  будет компонентой  $N_A$ -мерного марковского процесса, где

$N_A$  — удвоенное число степеней свободы четырехполюсника (возможно, и нечетное). Пусть теперь  $\mathcal{E}(t)$  не дельта-коррелирована, так что ток  $I(t)$  не является компонентой  $N_A$ -мерного марковского процесса. Однако если можно подобрать такой че-

тырехполюсник  $B$ , который под действием дельта-коррелированного шума  $e(t)$  дает на выходе э. д. с.  $\mathcal{E}(t)$ , то тот же ток  $I(t)$  будет откликом на дельта-коррелированную силу в системе  $A+B$  (рис. 27, б). Таким образом,  $I(t)$  будет компонентой  $(N_B + N_A)$ -мерного марковского процесса.

Возьмем для примера уже рассмотренную  $RL$ -ячейку, и пусть дельта-коррелированная э. д. с.  $e(t)$  действует на нее через фильтр  $L_0, C_0, R_0$  (рис. 28). Состояние фильтра описывается, скажем, зарядом  $q$  и током  $\dot{q}$  (1 степень свободы), а состояние  $RL$ -ячейки — током  $I$  (1/2 степени свободы). Ток  $I$  будет компонентой трехмерного марковского процесса с вероятностью перехода  $v(t, q, \dot{q}, I | t_0, q_0, \dot{q}_0, I_0)$ . Если же дельта-коррелированная э. д. с. действует на  $RL$ -ячейку непосредственно, то ток  $I$  будет одномерным марковским процессом с вероятностью перехода  $v(t, I | t_0, I_0)$ .

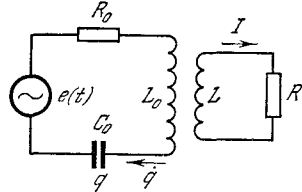


Рис. 28.

### Задачи

1. Рассчитать функцию корреляции тока в  $RL$ -контуре, на который действует стационарная случайная э. д. с.  $\mathcal{E}(t)$  с функцией корреляции

$$\psi_{\mathcal{E}}(\tau) = \frac{\alpha C}{2} e^{-\alpha |\tau|}.$$

Решение. Уравнение для тока  $I(t)$  есть

$$\dot{I} + \beta I = \mathcal{E}(t)/L, \quad \beta = R/L.$$

Импульсный отклик контура  $H(t) = e^{-\beta t}$  при  $t \geq 0$  и равен нулю при  $t < 0$ . Согласно (37.19) имеем

$$\begin{aligned} \psi_I(t, \tau) &= \int_0^t e^{-\beta \theta} d\theta \int_0^{t+\tau} e^{-\beta \theta'} \frac{\alpha C}{2L^2} e^{-\alpha |\tau + \theta - \theta'|} d\theta' = \\ &= \frac{\alpha C}{2L^2 (\alpha^2 - \beta^2)} \left\{ e^{-\beta \tau} \left[ \frac{\alpha}{\beta} (1 - e^{-2\beta t}) - e^{-2\beta t} + e^{-(\alpha + \beta)t} \right] - \right. \\ &\quad \left. - e^{-\alpha \tau} [1 - e^{-(\alpha + \beta)t}] \right\}. \quad (1) \end{aligned}$$

В частности, при  $\tau = 0$  получаем отсюда дисперсию тока:

$$D[I(t)] = \frac{\alpha C}{2L^2 (\alpha^2 - \beta^2)} \left\{ \frac{\alpha}{\beta} (1 - e^{-2\beta t}) - e^{-2\beta t} + 2e^{-(\alpha + \beta)t} - 1 \right\}. \quad (2)$$

Обе формулы справедливы при произвольном соотношении между временем релаксации контура  $(1/\beta)$  и временем корреляции случайной э. д. с.  $(1/\alpha)$ .

При  $\beta t \gg 1$  достигаются стационарные значения  $\psi_I$  и  $D[I]$ :

$$\psi_I(\tau) = \frac{\alpha C}{2L^2 (\alpha^2 - \beta^2)} \left( \frac{\alpha}{\beta} e^{-\beta \tau} - e^{-\alpha \tau} \right), \quad D[I] = \frac{\alpha C}{2L^2 \beta (\alpha + \beta)}.$$

При  $\alpha \rightarrow \infty$  мы приходим к дельта-коррелированной э. д. с.  $\mathcal{E}(t)$  и марковскому процессу  $I(t)$ . Формулы (1) и (2) принимают вид

$$\Psi_I(t, \tau) = \frac{C e^{-\beta \tau}}{2L^2 \beta} (1 - e^{-2\beta t}), \quad D[I(t)] = \frac{C}{2L^2 \beta} (1 - e^{-2\beta t}).$$

Формула для  $D[I]$  совпадает, конечно, с (35.11).

Если  $\beta \rightarrow 0$ , то, согласно (2),

$$D[I(t)] = \frac{C}{L^2 \alpha} (\alpha t + e^{-\alpha t} - 1),$$

что дает при  $\alpha t \gg 1$  диффузионный закон  $D[I] = \frac{C}{L^2} t$ , т. е. ток  $I(t) =$

$$= \frac{1}{L} \int_0^t \mathcal{E}(t) dt \text{ начинает вести себя как функция с независимыми прираще-}$$

ниями.

2. Используя результаты § 28, показать, что огибающая отклика селективного колебательного контура на белый шум удовлетворяет линейному стохастическому дифференциальному уравнению первого порядка.

Решение. Селективный контур — томсоновская система, описываемая уравнением

$$\frac{d^2 x}{dt'^2} + \mu h \frac{dx}{dt'} + x = \mu F(t'),$$

где  $t' = \omega_0 t$ ,  $\mu$  — малый параметр. По условию

$$\langle F(t'_1) F(t'_2) \rangle = C \delta(t'_1 - t'_2).$$

Уравнение контура отличается от уравнения (28.1) для томсоновского автогенератора лишь отсутствием нелинейного члена и знаком члена с  $dx/dt'$  [ $\mu h$  — декремент, в отличие от инкремента  $\mu p$  в (28.1)]. Следовательно, для огибающей  $A(\theta)$  отклика  $x = A(\theta) \cos[t' + \varphi(\theta)]$  мы получаем первое уравнение Ван-дер-Поля (28.7) с  $R(A) = -\frac{h}{2} A$ , т. е.

$$\frac{dA}{d\theta} + \frac{h}{2} A = -\frac{1}{2} F_{\perp}(\theta), \quad (1)$$

где  $\theta = \mu t'$  — медленное безразмерное время.

Если компонента силы  $F_{\perp}(\theta)$  дельта-коррелирована, то по общей теореме (§ 36)  $A(\theta)$  — диффузионный марковский процесс, функция корреляции которого, в соответствии с (1), экспоненциальна. То, что для селективной системы можно в первом приближении считать  $F_{\perp}(\theta)$  дельта-коррелированным процессом, показано далее, в § 53.

Теорема о марковости огибающей была высказана в таком виде (и довольно громоздко доказана) Пирсом [12]. На ее доказательство при помощи уравнения Смолуховского (или уравнения Эйнштейна — Фоккера) указал Хелстром [3]. Из последующих работ по статистическим свойствам огибающей упомянем [14, 15].

3. Написать уравнение Эйнштейна — Фоккера для осциллятора с флуктуирующей собственной частотой

$$\ddot{x} + 2h\dot{x} + \omega_0^2 [1 + f(t)] x = 0, \quad (1)$$

где  $f(t)$  — нормальный дельта-коррелированный случайный процесс:

$$\langle f(t) \rangle = 0, \quad \langle f(t) f(t') \rangle = 2\mathcal{D} \delta(t - t'), \quad \mathcal{D} = \text{const.}$$

Решение. Условия общей теоремы, приведенной в § 36, выполнены, и, следовательно, в осцилляторе — системе с одной степенью свободы — происходит двумерный диффузионный марковский процесс. Записав стохастическое уравнение (1) в виде двух уравнений первого порядка:

$$\begin{aligned}\dot{x} &= y \equiv X(x, y, t) + \dot{x}(x, y, t), \\ \dot{y} &= -2hy - \omega_0^2 [1 + f(t)] x \equiv Y(x, y, t) + \dot{y}(x, y, t),\end{aligned}\quad (2)$$

получаем

$$X = y, \quad Y = -2hy - \omega_0^2 x, \quad \dot{x} = 0, \quad \dot{y} = -\omega_0^2 x f(t).$$

Таким образом, в корреляционной матрице случайных воздействий  $\dot{x}$  и  $\dot{y}$  отличен от нуля только элемент

$$\langle \dot{y}_y(x, y, t; x', y', t') \rangle = \langle \dot{y}_y(x, y, t) \dot{y}_y(x', y', t') \rangle = 2\mathcal{D}\omega_0^2 x x' \delta(t - t').$$

В результате уравнение Эйнштейна — Фоккера (36.15) для плотности вероятности перехода  $v(x, y, t | x_0, y_0, 0)$  имеет вид

$$\frac{\partial v}{\partial t} = -y \frac{\partial v}{\partial x} + \frac{\partial (2hy + \omega_0^2 x) v}{\partial y} + \mathcal{D}\omega_0^4 x^2 \frac{\partial^2 v}{\partial y^2}. \quad (3)$$

Из-за квадратичной зависимости коэффициента диффузии от  $x$  это уравнение не может быть решено в общем виде.

4. Получить при помощи уравнения (3) предыдущей задачи уравнения для моментов  $M_{rs}(t) = \langle x^r(t) y^s(t) \rangle$ . Показать, что эти уравнения распадаются на замкнутые системы для моментов каждого фиксированного порядка  $r + s = n$ . Исследовать систему уравнений для моментов второго порядка

$$M_{20} = \overline{x^2}, \quad M_{11} = \overline{xy}, \quad M_{02} = \overline{y^2}$$

и найти условие параметрического возбуждения осциллятора в данном случае флуктуаций его частоты в виде нормального белого шума.

Решение. Умножим уравнение (3) на  $x^r y^s$  и проинтегрируем по  $x$  и  $y$  от  $-\infty$  до  $+\infty$ . Учитывая, что

$$M_{rs}(t) = \iint_{-\infty}^{+\infty} x^r y^s v(x, y, t | x_0, y_0, 0) dx dy,$$

и предполагая, что проинтегрированные члены обращаются в нуль при  $|x| = \infty$  или  $|y| = \infty$ , получаем для  $M_{rs}(t)$  следующую систему уравнений:

$$\dot{M}_{rs} = r M_{r-1, s+1} - s(2h M_{rs} + \omega_0^2 M_{r+1, s-1}) - \mathcal{D}\omega_0^4 s(s-1) M_{r+2, s-2}, \quad (4)$$

которую надо решать с начальными условиями  $M_{rs}(0) = x_0^r y_0^s$ .

Сумма индексов при всех  $M$  в (4) одинакова (равна  $r + s = n$ ), откуда и следует, что для моментов порядка  $n$  уравнения (4) образуют замкнутую систему из  $n + 1$  уравнений. Полагая в (4)  $r = 1, s = 0$  и  $r = 0, s = 1$ , получаем уравнения для моментов первого порядка  $M_{10} = \bar{x}$  и  $M_{01} = \bar{y}$ :

$$\dot{\bar{x}} = \bar{y}, \quad \dot{\bar{y}} = -(2h\bar{y} + \omega_0^2 \bar{x}),$$

которые вытекают и из исходных уравнений (2) в отсутствие флуктуаций частоты [ $f(t) = 0$ ]. Таким образом,  $\bar{x}$  и  $\bar{y}$  меняются по динамическому закону.

Для моментов второго порядка получаем из (4) систему трех уравнений:

$$\begin{aligned}\dot{M}_{20} &= 2M_{11}, \quad \dot{M}_{11} = M_{02} - (2hM_{11} + \omega_0^2 M_{20}), \\ \dot{M}_{02} &= -2(2hM_{02} + \omega_0^2 M_{11}) + 2\mathcal{D}\omega_0^4 M_{20}.\end{aligned}$$

Ищем частное решение этих обыкновенных линейных уравнений с постоянными коэффициентами в виде

$$M_{20} = C_1 e^{\lambda t}, \quad M_{11} = C_2 e^{\lambda t}, \quad M_{02} = C_3 e^{\lambda t},$$

что дает для  $C_1$ ,  $C_2$  и  $C_3$  систему однородных алгебраических уравнений

$$\begin{aligned}\lambda C_1 &= 2C_2, \quad \lambda C_2 = C_3 - (2hC_2 + \omega_0^2 C_1), \\ \lambda C_3 &= -2(2hC_3 + \omega_0^2 C_2) + 2\mathcal{D}\omega_0^4 C_1.\end{aligned}$$

Для того чтобы эта система имела нетривиальное решение, надо приравнять нулю ее определитель, что дает характеристическое уравнение для  $\lambda$ :

$$\lambda^3 + 6h\lambda^2 + 4(\omega_0^2 + 2h^2)\lambda + 4\omega_0^2(2h - \omega_0^2\mathcal{D}) = 0. \quad (5)$$

Как известно, необходимые и достаточные условия устойчивости, т. е. отсутствия положительных вещественных частей у корней  $\lambda_k$  уравнения (5), — это условия Рауса — Гурвица, которые для кубического уравнения

$$\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 = 0$$

имеют вид

$$a_1 > 0, \quad a_2 > 0, \quad a_3 > 0, \quad a_1 a_2 - a_3 > 0.$$

Для уравнения (5) эти условия таковы:

$$h > 0, \quad \omega_0^2 + 2h^2 > 0, \quad 2h - \omega_0^2\mathcal{D} > 0, \quad 4h(\omega_0^2 + 3h^2) + \omega_0^4\mathcal{D} > 0.$$

Нарушаться может только третье неравенство, т. е. при

$$\mathcal{D} > 2h/\omega_0^2 \quad (6)$$

и малейшем отклонении начальных значений от нуля будет происходить экспоненциальное нарастание моментов  $\overline{x^2}$ ,  $\overline{xy}$  и  $\overline{y^2}$  и, в частности, средней энергии осциллятора (параметрическое возбуждение).

Исследование уравнений для моментов порядка  $n > 2$  показывает, что условия параметрического возбуждения для разных  $n$  различны. В частности, для четвертых моментов это условие слабее (6), а именно:

$$\mathcal{D} > \frac{2h}{3\omega_0^2} \frac{\omega_0^2 + 3h^2}{\omega_0^2 + 6h^2}.$$

Те же замкнутые системы уравнений (4) для моментов можно получить и непосредственно из стохастических уравнений (2), не предполагая нормальности  $f(t)$ , но считая, что высшие моменты  $f(t)$  выражаются через произведения дельта-функций  $\delta(t_2 - t_1)\delta(t_3 - t_1)\dots\delta(t_n - t_1)$ . Исследование условий применимости диффузионного приближения [7] показывает, что при наличии у  $f(t)$  не равного нулю времени корреляции  $\theta_f$  должны выполняться условия

$$\theta_f \ll T_0 = 2\pi/\omega_0, \quad \theta_f \ll 1/\omega_0^2\mathcal{D}, \quad t \gg \theta_f.$$

## Глава I

### ОБЩИЕ СВЕДЕНИЯ О СЛУЧАЙНЫХ ПОЛЯХ

#### § 1. Основные понятия

В части I этой книги мы, почти без исключений, имели дело с однопараметрическими случайными функциями  $\xi(\alpha)$ , причем в большинстве приложений параметр  $\alpha$  являлся временем  $t$ . Если же речь идет о случайной функции более чем одного параметра,  $\xi(\alpha, \beta, \dots)$ , то ее называют обычно *случайным полем* (в пространстве параметров  $\alpha, \beta, \dots$ ). Мы сразу же ограничимся случаем, когда параметров всего четыре: время  $t$  и точка пространства  $\mathbf{r} = (x, y, z)$ . Можно, конечно, называть  $\xi(t, \mathbf{r})$  полем в четырехмерном пространстве, но в нерелятивистских задачах привычнее и нагляднее говорить о переменном (зависящем от  $t$ ) поле в трехмерном пространстве  $(x, y, z)$ .

В свою очередь случайное поле может описываться не одной, а  $N$  функциями  $\xi^{(i)}(t, \mathbf{r})$ ,  $i = 1, 2, \dots, N$ , и называется тогда  $N$ -мерным случайным полем, подобно  $N$ -мерной случайной функции  $\xi^{(i)}(t)$ ,  $i = 1, 2, \dots, N$ . С чисто математической точки зрения компоненты  $\xi^{(i)}(t, \mathbf{r})$  могут быть чем угодно и даже не обязаны обладать одинаковой размерностью. Например, флуктуации плотности жидкости  $\rho$ , давления  $p$ , температуры  $T$  и скорости  $\mathbf{v}$  образуют в совокупности шестимерное случайное поле. Но особый физический интерес представляют, конечно, те случаи, когда величины  $\xi^{(i)}(t, \mathbf{r})$  одномерны и обладают определенными трансформационными свойствами при ортогональных преобразованиях координат в пространстве  $x, y, z$ , т. е. являются совокупностью компонент тензора какого-либо ранга. При таком подходе целесообразнее говорить в приведенном примере о *четырех* полях — трех скалярных ( $\rho, p$  и  $T$ ) и одном векторном ( $\mathbf{v}$ ).

Ряд определений и свойств, введенных и установленных ранее в ч. I для случайных функций одного параметра, естественным образом обобщается и на случайные поля, зависящие

от многих параметров, в частности, на пространственно-временные случайные поля, зависящие от  $x, y, z, t$ . Эти (зачастую довольно очевидные) обобщения касаются и вопроса о том, что означает *полное задание* случайного поля.

Обозначим для краткости через  $Q = (t, r)$  точку в четырехмерном мире. *Полное задание* одномерного случайного поля  $\xi(Q)$  означает, что известны все его  $n$ -мерные, или, как иногда говорят,  $n$ -точечные ( $n = 1, 2, \dots$ ), плотности вероятностей, т. е. для любого числа  $n$  произвольно выбранных точек  $Q_v$  известны функции

$$\omega_n \{ \xi_1, \dots, \xi_n \} d\xi_1 \dots d\xi_n = \\ = P \{ \xi_v \leq \xi(Q_v) < \xi_v + d\xi_v \quad (v = 1, 2, \dots, n) \}, \quad (1.1)$$

где  $P \{ \xi \leq \xi(Q) < \xi + d\xi \}$  — вероятность того, что случайная величина  $\xi(Q)$  приняла значение, лежащее в интервале  $(\xi, \xi + d\xi)^1$ . Аналогично, полное статистическое задание (описание)  $N$ -мерного поля  $\xi^{(i)}(Q)$  дается совокупностью  $nN$ -мерных плотностей вероятностей

$$\omega_{nN} \{ \xi_1^{(1)}, \xi_1^{(2)}, \dots, \xi_n^{(N)} \} d\xi_1^{(1)} d\xi_1^{(2)} \dots d\xi_n^{(N)} = \\ = P \{ \xi_v^{(i)} \leq \xi^{(i)}(Q_v) < \xi_v^{(i)} + d\xi_v^{(i)}, v = 1, 2, \dots, n, i = 1, 2, \dots, N \}. \quad (1.2)$$

Эти плотности вероятностей, разумеется, должны быть подчинены условиям неотрицательности, симметрии, согласованности и нормировки (ч. I, § 14).

Очевидным образом распространяется на случайные поля и понятие статистической однородности. Одномерное случайное поле  $\xi$  называется *однородным* (в узком смысле), т. е. стационарным<sup>2)</sup> по  $t$  и однородным по  $x, y, z$ , если все  $n$ -мерные плотности вероятности  $\omega_n$  инвариантны относительно преобразования трансляции  $Q \rightarrow Q + \delta Q$ :

$$P \{ \xi_v \leq \xi(Q_v + \delta Q) < \xi_v + d\xi_v, v = 1, 2, \dots, n \} = \\ = P \{ \xi_v \leq \xi(Q_v) < \xi_v + d\xi_v, v = 1, 2, \dots, n \}. \quad (1.3)$$

Если речь идет о  $N$ -мерном случайном поле  $\xi^{(i)}$  и указанной инвариантностью относительно сдвига  $\delta Q$  обладают все  $nN$ -мерные плотности вероятностей  $\omega_{nN}$ , то говорят об *однородных и однородно связанных* (в  $Q$ -пространстве) полях  $\xi^{(i)}$  ( $i = 1, 2, \dots, N$ ).

Понятие *пространственной однородности* для случайных полей (инвариантность плотностей  $\omega_n$  по отношению к простран-

<sup>1)</sup> Полное описание случайного поля достигается также при помощи характеристических функционалов, о которых пойдет речь в § 7.

<sup>2)</sup> Однородность по времени принято называть стационарностью.



ственному сдвигу  $r \rightarrow r + \delta r$ ) является естественным обобщением понятия стационарности для случайных функций времени (ч. I, § 16). Но многомерность пространства параметров  $x, y, z, t$  открывает новые возможности, а именно: поле может быть по части параметров однородным, а по остальным — неоднородным. Например, наряду с полями, однородными в  $Q$ -пространстве, иногда приходится иметь дело со стационарными, но пространственно-неоднородными полями или же однородными, но нестационарными полями. В волновых задачах часто встречаются поля, однородные только на определенных поверхностях, скажем на плоскости или на сфере.

Зная многомерные плотности вероятностей, можно вычислить моменты случайного поля  $\xi$  любого порядка. В общем случае эти моменты будут функциями координат:  $Q_v = (t_v, r_v)$ . Во многих вопросах наибольший интерес представляют наинизшие моменты (первого и второго порядка), с которыми оперирует корреляционная теория случайных полей. Основные понятия этой теории те же, что и в корреляционной теории случайных функций.

*Среднее значение* случайного поля  $\xi(Q)$  (момент первого порядка) вычисляется при помощи одномерной плотности вероятностей  $\omega_1(\xi)$ <sup>1)</sup>:

$$\langle \xi(Q) \rangle = \int \xi \omega_1(\xi) d\xi. \quad (1.4)$$

Флуктуационную часть случайного поля  $\xi$ , т. е. величину  $\xi - \langle \xi \rangle$ , мы по-прежнему будем обозначать волнистой чертой сверху:

$$\tilde{\xi} \equiv \xi - \langle \xi \rangle \equiv \xi - \bar{\xi}. \quad (1.5)$$

*Смешанный момент* второго порядка  $B_{\xi}$  вычисляется при помощи двумерной плотности вероятностей  $\omega_2(\xi_1, \xi_2)$ :

$$B_{\xi}(Q_1, Q_2) \equiv \langle \xi(Q_1) \xi(Q_2) \rangle = \iint \xi_1 \xi_2 \omega_2(\xi_1, \xi_2) d\xi_1 d\xi_2. \quad (1.6)$$

Через нее же выражается и *функция корреляции* (точнее, функция автокорреляции):

$$\begin{aligned} \psi_{\xi}(Q_1, Q_2) &\equiv \langle \tilde{\xi}(Q_1) \tilde{\xi}(Q_2) \rangle = \\ &= \iint (\xi_1 - \langle \xi_1 \rangle)(\xi_2 - \langle \xi_2 \rangle) \omega_2(\xi_1, \xi_2) d\xi_1 d\xi_2 = \\ &= B_{\xi}(Q_1, Q_2) - \langle \xi(Q_1) \rangle \langle \xi(Q_2) \rangle. \end{aligned} \quad (1.7)$$

Для случайных полей с нулевым средним значением функции  $B_{\xi}(Q_1, Q_2) = \langle \xi(Q_1) \xi(Q_2) \rangle$  и  $\psi_{\xi}(Q_1, Q_2)$  совпадают. Дисперсия

<sup>1)</sup> Операция статистического усреднения обозначена здесь посредством угловых скобок, но иногда, если это будет удобнее, мы будем пользоваться, как и в ч. I, прямой чертой сверху, так что  $\langle \xi \rangle \equiv \bar{\xi}$ .

случайного поля  $D[\xi]$ , т. е. средний квадрат флуктуаций, равен

$$D[\xi] = \sigma_{\xi}^2(Q) = \langle \bar{\xi}^2(Q) \rangle = \langle [\xi(Q) - \langle \xi(Q) \rangle]^2 \rangle = \psi_{\xi}(Q, Q). \quad (1.8)$$

$\xi^{(i)}$ . Многомерное поле  $\xi^{(i)}(Q)$  в рамках корреляционной теории характеризуется совокупностью средних значений  $\langle \xi^{(i)}(Q) \rangle$  и матрицей моментов второго порядка с элементами

$$B_{ik}(1, 2) = \langle \xi^{(i)}(1) \xi^{(k)}(2) \rangle \quad (1.9)$$

или, что равносильно, — *корреляционной матрицей* с элементами

$$\psi_{ik}(1, 2) = \langle \bar{\xi}^{(i)}(1) \bar{\xi}^{(k)}(2) \rangle = B_{ik}(1, 2) - \langle \xi^{(i)}(1) \rangle \langle \xi^{(k)}(2) \rangle. \quad (1.10)$$

Для краткости аргументы  $Q_1$  и  $Q_2$  здесь заменены просто своими номерами 1 и 2. Диагональные элементы корреляционной матрицы представляют собой *функции автокорреляции*  $\psi_{kk}(1, 2)$ , а недиагональные — *функции взаимной корреляции*  $\psi_{ik}(1, 2)$  полей  $\xi^{(i)}$  и  $\xi^{(k)}$ .

Наряду с вещественными случайными полями  $\xi(Q)$ , о которых шла речь выше, часто приходится рассматривать также *комплексные поля*

$$\zeta(Q) = \xi(Q) + i\eta(Q), \quad (1.11)$$

где  $\xi = \text{Re } \zeta$  и  $\eta = \text{Im } \zeta$  — вещественные функции в  $Q$ -пространстве. Полное статистическое описание комплексного случайного поля  $\zeta(Q)$  осуществляется заданием  $2n$ -мерных плотностей вероятностей  $\omega_{2n}(\xi_1, \dots, \xi_n; \eta_1, \dots, \eta_n)$  ( $n = 1, 2, \dots$ ), через которые выражаются вероятности

$$P\{\xi_{\mu} \leq \xi(Q_{\mu}) < \xi_{\mu} + d\xi_{\mu}, \eta_{\nu} \leq \eta(Q_{\nu}) < \eta_{\nu} + d\eta_{\nu} \\ (\mu, \nu = 1, 2, \dots, n)\}$$

аналогично (1.1). Через  $\omega_{2n}$  можно выразить любые моменты комплексного случайного поля  $\zeta$  и, в частности, его низшие моменты. Последние представляют для рассматриваемых далее задач наибольший интерес, так что мы лишь изредка будем выходить за пределы корреляционной теории случайных полей.

## § 2. Пространственные корреляционные функции комплексных случайных полей

Рассмотрим сначала случайные поля, зависящие только от пространственных координат, т. е. не зависящие от времени. К полям, зависящим также от времени, мы обратимся в § 6.

Итак, пусть одномерное комплексное случайное поле зависит только от  $r$ :

$$\zeta(r) = \xi(r) + i\eta(r).$$

По определению корреляционная (*автокорреляционная*) функция комплексного поля  $\xi$  равна

$$\psi_{\xi}(\mathbf{r}_1, \mathbf{r}_2) \equiv \langle \tilde{\xi}(\mathbf{r}_1) \tilde{\xi}^*(\mathbf{r}_2) \rangle = \langle \xi(\mathbf{r}_1) \xi^*(\mathbf{r}_2) \rangle - \langle \xi(\mathbf{r}_1) \rangle \langle \xi^*(\mathbf{r}_2) \rangle. \quad (2.1)$$

Положив здесь  $\mathbf{r}_1 = \mathbf{r}_2 = \mathbf{r}$ , получим дисперсию в точке  $\mathbf{r}$ :

$$\sigma_{\xi}^2(\mathbf{r}) \equiv \langle |\tilde{\xi}(\mathbf{r})|^2 \rangle = \sigma_{\xi}^2 + \sigma_{\eta}^2 = \psi_{\xi}(\mathbf{r}, \mathbf{r}). \quad (2.2)$$

Многие свойства пространственной корреляционной функции  $\psi_{\xi}(\mathbf{r}_1, \mathbf{r}_2)$  имеют аналоги в теории случайных процессов (ч. I, §§ 14, 17, 18 и 38). Перечислим кратко эти свойства.

Корреляционная функция  $\psi_{\xi}(\mathbf{r}_1, \mathbf{r}_2)$  является *эрмитовой*:

$$\psi_{\xi}(\mathbf{r}_1, \mathbf{r}_2) = \psi_{\xi}^*(\mathbf{r}_2, \mathbf{r}_1), \quad (2.3)$$

как это непосредственно следует из определения (2.1). В частном случае *вещественного* случайного поля ( $\xi = \xi^*$ ) корреляционная функция *симметрична*:

$$\psi_{\xi}(\mathbf{r}_1, \mathbf{r}_2) = \psi_{\xi}(\mathbf{r}_2, \mathbf{r}_1). \quad (2.4)$$

Далее, квадрат модуля  $\psi_{\xi}(\mathbf{r}_1, \mathbf{r}_2)$  никогда не превышает произведения дисперсий  $\sigma_{\xi}^2(\mathbf{r}_1)$  и  $\sigma_{\xi}^2(\mathbf{r}_2)$ :

$$|\psi_{\xi}(\mathbf{r}_1, \mathbf{r}_2)|^2 \leq \sigma_{\xi}^2(\mathbf{r}_1) \sigma_{\xi}^2(\mathbf{r}_2) = \psi_{\xi}(\mathbf{r}_1, \mathbf{r}_1) \psi_{\xi}(\mathbf{r}_2, \mathbf{r}_2). \quad (2.5)$$

Из (2.5) вытекает, в частности, что *коэффициент корреляции* (иначе — *нормированная корреляционная функция*), определяемый как

$$K_{\xi}(\mathbf{r}_1, \mathbf{r}_2) = \frac{\psi_{\xi}(\mathbf{r}_1, \mathbf{r}_2)}{\sigma_{\xi}(\mathbf{r}_1) \sigma_{\xi}(\mathbf{r}_2)} = \frac{\psi_{\xi}(\mathbf{r}_1, \mathbf{r}_2)}{\sqrt{\psi_{\xi}(\mathbf{r}_1, \mathbf{r}_1) \psi_{\xi}(\mathbf{r}_2, \mathbf{r}_2)}}, \quad (2.6)$$

не превышает по модулю единицы:

$$|K_{\xi}(\mathbf{r}_1, \mathbf{r}_2)| \leq 1. \quad (2.7)$$

Из неотрицательности величины  $\langle |\int \tilde{\xi}(\mathbf{r}) u(\mathbf{r}) d^3r|^2 \rangle$  следует, что корреляционная функция  $\psi_{\xi}(\mathbf{r}_1, \mathbf{r}_2)$  является *положительно определенной*, т. е.

$$\iint_V \psi_{\xi}(\mathbf{r}_1, \mathbf{r}_2) u(\mathbf{r}_1) u^*(\mathbf{r}_2) d^3r_1 d^3r_2 \geq 0, \quad (2.8)$$

где  $u(\mathbf{r})$  — произвольная комплексная функция, а  $V$  — произвольная область интегрирования, для которых интеграл (2.8) существует.

*Взаимной корреляционной функцией* двух комплексных полей  $\mu(\mathbf{r})$  и  $\nu(\mathbf{r})$  называется величина

$$\psi_{\mu\nu}(\mathbf{r}_1, \mathbf{r}_2) = \langle \tilde{\mu}(\mathbf{r}_1) \tilde{\nu}^*(\mathbf{r}_2) \rangle. \quad (2.9)$$

Квадрат модуля  $\psi_{\mu\nu}$  не превышает произведения дисперсий  $\sigma_\mu^2$  и  $\sigma_\nu^2$ :

$$|\psi_{\mu\nu}(\mathbf{r}_1, \mathbf{r}_2)|^2 \leq \sigma_\mu^2(\mathbf{r}_1) \sigma_\nu^2(\mathbf{r}_2), \quad (2.10)$$

что аналогично (2.5). Как следствие этого, *взаимный коэффициент корреляции*

$$K_{\mu\nu}(\mathbf{r}_1, \mathbf{r}_2) = \frac{\psi_{\mu\nu}(\mathbf{r}_1, \mathbf{r}_2)}{\sigma_\mu(\mathbf{r}_1) \sigma_\nu(\mathbf{r}_2)} \quad (2.11)$$

по модулю не превышает единицы:  $|K_{\mu\nu}| \leq 1$ . Существенно, что, в отличие от автокорреляционных функций, взаимная функция корреляции  $\psi_{\mu\nu}$  не обладает ни свойством эрмитовости (2.3), ни положительной определенностью (2.8), но при этом  $\psi_{\mu\nu}^*(\mathbf{r}_1, \mathbf{r}_2) = \psi_{\nu\mu}(\mathbf{r}_2, \mathbf{r}_1)$ .

Наряду с  $\psi_\xi(\mathbf{r}_1, \mathbf{r}_2)$  иногда полезно ввести еще одну автокорреляционную функцию комплексного поля:

$$\tilde{\psi}_\xi(\mathbf{r}_1, \mathbf{r}_2) = \langle \tilde{\xi}(\mathbf{r}_1) \tilde{\xi}(\mathbf{r}_2) \rangle, \quad (2.12)$$

которая ранее (ч. I, § 38) была названа *второй* корреляционной функцией. В отличие от  $\psi_\xi(\mathbf{r}_1, \mathbf{r}_2)$ , т. е. от *первой* функции корреляции, определенной выражением (2.1), значение случайного поля  $\tilde{\xi}(\mathbf{r}_2)$  входит в (2.12) без комплексного сопряжения. Вторая корреляционная функция симметрична относительно своих аргументов,  $\tilde{\psi}_\xi(\mathbf{r}_2, \mathbf{r}_1) = \tilde{\psi}_\xi(\mathbf{r}_1, \mathbf{r}_2)$ , а для вещественного поля  $\xi$  она совпадает, очевидно, с  $\psi_\xi(\mathbf{r}_1, \mathbf{r}_2)$ .

Через первую и вторую функции корреляции  $\psi_\xi(\mathbf{r}_1, \mathbf{r}_2)$  и  $\tilde{\psi}_\xi(\mathbf{r}_1, \mathbf{r}_2)$  можно выразить автокорреляционные функции вещественной и мнимой частей комплексного поля  $\zeta = \xi + i\eta$ :

$$\psi_\xi(\mathbf{r}_1, \mathbf{r}_2) = \langle \tilde{\xi}(\mathbf{r}_1) \tilde{\xi}(\mathbf{r}_2) \rangle, \quad \psi_\eta(\mathbf{r}_1, \mathbf{r}_2) = \langle \tilde{\eta}(\mathbf{r}_1) \tilde{\eta}(\mathbf{r}_2) \rangle,$$

а также их взаимную функцию корреляции:

$$\psi_{\xi\eta}(\mathbf{r}_1, \mathbf{r}_2) = \langle \tilde{\xi}(\mathbf{r}_1) \tilde{\eta}(\mathbf{r}_2) \rangle = \psi_{\eta\xi}(\mathbf{r}_2, \mathbf{r}_1).$$

Заметим, что симметрии относительно аргументов  $\mathbf{r}_1$  и  $\mathbf{r}_2$  у функции  $\psi_{\xi\eta}$ , вообще говоря, нет. Используя очевидные соотношения

$$\begin{aligned} \psi_\xi(\mathbf{r}_1, \mathbf{r}_2) &= \psi_\xi(\mathbf{r}_2, \mathbf{r}_1) + \psi_\eta(\mathbf{r}_1, \mathbf{r}_2) - i\psi_{\xi\eta}(\mathbf{r}_1, \mathbf{r}_2) + i\psi_{\eta\xi}(\mathbf{r}_1, \mathbf{r}_2), \\ \tilde{\psi}_\xi(\mathbf{r}_1, \mathbf{r}_2) &= \psi_\xi(\mathbf{r}_1, \mathbf{r}_2) - \psi_\eta(\mathbf{r}_1, \mathbf{r}_2) + i\psi_{\xi\eta}(\mathbf{r}_1, \mathbf{r}_2) + i\psi_{\eta\xi}(\mathbf{r}_1, \mathbf{r}_2), \end{aligned} \quad (2.13)$$

получаем следующие выражения для компонент корреляционной матрицы вещественных полей  $\xi(\mathbf{r})$  и  $\eta(\mathbf{r})$ :

$$\begin{aligned}\psi_{\xi}(\mathbf{r}_1, \mathbf{r}_2) &= 1/2 \operatorname{Re} [\psi_{\xi}(\mathbf{r}_1, \mathbf{r}_2) + \bar{\psi}_{\xi}(\mathbf{r}_1, \mathbf{r}_2)], \\ \psi_{\eta}(\mathbf{r}_1, \mathbf{r}_2) &= 1/2 \operatorname{Re} [\psi_{\xi}(\mathbf{r}_1, \mathbf{r}_2) - \bar{\psi}_{\xi}(\mathbf{r}_1, \mathbf{r}_2)], \\ \psi_{\xi\eta}(\mathbf{r}_1, \mathbf{r}_2) &= 1/2 \operatorname{Im} [\bar{\psi}_{\xi}(\mathbf{r}_1, \mathbf{r}_2) - \psi_{\xi}(\mathbf{r}_1, \mathbf{r}_2)], \\ \psi_{\eta\xi}(\mathbf{r}_1, \mathbf{r}_2) &= 1/2 \operatorname{Im} [\bar{\psi}_{\xi}(\mathbf{r}_1, \mathbf{r}_2) + \psi_{\xi}(\mathbf{r}_1, \mathbf{r}_2)].\end{aligned}\quad (2.14)$$

От аналогичных формул (38.3) в ч. I формулы (2.14) отличаются только тем, что вместо моментов времени  $t$  и  $t'$  в них фигурируют радиусы-векторы  $\mathbf{r}_1$  и  $\mathbf{r}_2$ .

В случае статистически однородного (в широком смысле) случайного поля  $\xi$  инвариантность относительно сдвига  $\mathbf{r} \rightarrow \mathbf{r} + \delta\mathbf{r}$  должна выполняться только для среднего значения и для моментов второго порядка. Иными словами, должно быть

$$\langle \xi(\mathbf{r}) \rangle = \langle \xi(\mathbf{r} + \delta\mathbf{r}) \rangle, \quad (2.15)$$

$$\psi_{\xi}(\mathbf{r}_1, \mathbf{r}_2) = \psi_{\xi}(\mathbf{r}_1 + \delta\mathbf{r}, \mathbf{r}_2 + \delta\mathbf{r}). \quad (2.16)$$

Условие (2.15) означает, что  $\langle \xi \rangle = \text{const}$ , т. е. среднее значение является постоянной величиной. Положив в (2.16)  $\delta\mathbf{r} = -\mathbf{r}_2$ , находим, что  $\psi_{\xi}(\mathbf{r}_1, \mathbf{r}_2) = \psi_{\xi}(\mathbf{r}_1 - \mathbf{r}_2, 0)$ , т. е. корреляционная функция  $\psi_{\xi}$  статистически однородного поля зависит только от разности  $\mathbf{r}_1 - \mathbf{r}_2$ , но не от  $\mathbf{r}_1$  и  $\mathbf{r}_2$  порознь. Для краткости вместо  $\psi_{\xi}(\mathbf{r}_1 - \mathbf{r}_2, 0)$  принято писать просто  $\psi_{\xi}(\mathbf{r})$ :

$$\psi_{\xi}(\mathbf{r}) = \langle \bar{\xi}(\mathbf{r}_1) \xi^*(\mathbf{r}_1 - \mathbf{r}) \rangle. \quad (2.17)$$

Подобным же образом (только без комплексного сопряжения) записывается и вторая корреляционная функция  $\bar{\psi}_{\xi}(\mathbf{r})$ .

Дисперсия статистически однородного поля постоянна:

$$\sigma_{\xi}^2 \equiv \langle |\bar{\xi}(\mathbf{r})|^2 \rangle = \psi_{\xi}(0), \quad (2.18)$$

а общие свойства (2.3), (2.5) и (2.8) корреляционной функции для статистически однородных полей принимают следующую форму, аналогичную соотношениям для стационарных случайных процессов:

$$\psi_{\xi}(\mathbf{r}) = \psi_{\xi}^*(-\mathbf{r}), \quad (2.19)$$

$$|\psi_{\xi}(\mathbf{r})| \leq \sigma_{\xi}^2. \quad (2.20)$$

$$\int_V d^3r_1 \int_V d^3r_2 \psi_{\xi}(\mathbf{r}_1 - \mathbf{r}_2) u(\mathbf{r}_1) u^*(\mathbf{r}_2) \geq 0. \quad (2.21)$$

Согласно (2.19)

$$\operatorname{Re} \psi_{\xi}(-\mathbf{r}) = \operatorname{Re} \psi_{\xi}(\mathbf{r}), \quad \operatorname{Im} \psi_{\xi}(-\mathbf{r}) = -\operatorname{Im} \psi_{\xi}(\mathbf{r}),$$

т. е. вещественная часть  $\psi_{\xi}$  четна по  $\mathbf{r}$ , а мнимая — нечетна. Четность функции  $\psi_{\xi}(\mathbf{r})$  автоматически влечет за собой ее вещественность.

Вторая корреляционная функция, симметричная по своим аргументам, для однородных полей становится четной:  $\bar{\psi}_{\xi}(-\mathbf{r}) = \bar{\psi}_{\xi}(\mathbf{r})$ . Для частного случая, когда четна не только вторая, но и первая корреляционная функция, выражения (2.14) принимают вид

$$\begin{aligned}\psi_{\xi}(\mathbf{r}) &= 1/2 [\psi_{\xi}(\mathbf{r}) + \operatorname{Re} \bar{\psi}_{\xi}(\mathbf{r})], \quad \psi_{\eta}(\mathbf{r}) = 1/2 [\psi_{\xi}(\mathbf{r}) - \operatorname{Re} \bar{\psi}_{\xi}(\mathbf{r})], \\ \psi_{\xi\eta}(\mathbf{r}) &= \psi_{\eta\xi}(\mathbf{r}) = 1/2 \operatorname{Im} \bar{\psi}_{\xi}(\mathbf{r}).\end{aligned}\quad (2.22)$$

Из этих соотношений следует, что вещественная и мнимая части  $\xi$  и  $\eta$  комплексного статистически однородного поля  $\zeta = \xi + i\eta$  тоже статистически однородны и, кроме того, однородно связаны.

Статистически однородные поля, у которых  $\psi_{\xi}(\mathbf{r})$  и  $\bar{\psi}_{\xi}(\mathbf{r})$  зависят *только от модуля* (но не от направления) вектора  $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ , соединяющего точки  $\mathbf{r}_1$  и  $\mathbf{r}_2$ ,

$$\psi_{\xi}(\mathbf{r}) = \psi_{\xi}(r), \quad \bar{\psi}_{\xi}(\mathbf{r}) = \bar{\psi}_{\xi}(r), \quad (2.23)$$

где  $r = |\mathbf{r}| = \sqrt{x^2 + y^2 + z^2}$ , называются *статистически изотропными*. Полиоправного аналога изотропных полей в теории случайных процессов указать нельзя. Ограниченную аналогию можно провести лишь со стационарными вещественными процессами, для которых функция корреляции четна:  $\psi(\tau) = \psi(|\tau|)$ .

Для изотропных случайных полей корреляционная функция всегда четна и, следовательно, всегда вещественна. Из (2.22) следует, что если поле  $\xi$  изотропно, то поля  $\xi = \operatorname{Re} \zeta$  и  $\eta = \operatorname{Im} \zeta$  изотропны и изотропно связаны между собой.

Примерами корреляционной функции однородного и изотропного случайного поля  $\zeta$  могут служить корреляционная функция в виде гауссовой кривой

$$\psi_{\xi}(r) = \sigma_{\xi}^2 e^{-r^2/l^2} \quad (2.24)$$

и экспоненциальная функция корреляции

$$\psi_{\xi}(r) = \sigma_{\xi}^2 e^{-r/l}. \quad (2.25)$$

В обоих примерах  $l$  — *радиус корреляции* поля, т. е. расстояние  $r = |\mathbf{r}_1 - \mathbf{r}_2|$ , на котором  $\psi_{\xi}(r)$  уменьшается примерно вдвое по сравнению с дисперсией  $\sigma_{\xi}^2$ . Для полей с функцией корреляции произвольного вида *эффективный* (или *интегральный*) радиус корреляции обычно определяется как

$$l_{\Phi\Phi} = \frac{1}{\sigma_{\xi}^2} \int_0^{\infty} \psi_{\xi}(r) dr = \int_0^{\infty} K_{\xi}(r) dr. \quad (2.26)$$

У гауссовой функции корреляции (2.24)  $l_{\phi\phi} = \sqrt{1/2} \pi l$ , а у экспоненциальной функции (2.25)  $l_{\phi\phi} = l$ .

Иногда встречаются корреляционные функции, имеющие определенный масштаб изменения, но для которых эффективного радиуса корреляции не существует из-за расходимости интеграла в (2.26). Так обстоит дело, например, для функции корреляции вида

$$\psi_{\zeta}(r) = \sigma_{\zeta}^2 (1 + r^2/l^2)^{-\mu}, \quad \mu < 1/2.$$

Хотя понятие радиуса корреляции (как и времени корреляции для случайных процессов) часто оказывается полезным, универсальное его определение, пригодное для флуктуаций произвольного вида, дать нельзя. Корреляционные функции часто обладают не одним, а несколькими характерными масштабами, как, например, быстро осциллирующая корреляционная функция с плавно меняющейся огибающей, или же корреляционные функции анизотропных полей, к которым мы и перейдем.

У статистически однородных, но анизотропных полей функции корреляции зависят не только от модуля, но и от направления вектора  $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ , как, скажем, в следующих двух примерах:

$$\psi_{\zeta}(\mathbf{r}) = \psi_{\zeta}[(\alpha x + \beta y + \gamma z)/l], \quad (2.27)$$

$$\psi_{\zeta}(\mathbf{r}) = \psi_{\zeta}\left(\frac{|x|}{a}, \frac{|y|}{b}, \frac{|z|}{c}\right). \quad (2.28)$$

Анизотропные поля, очевидно, не имеют аналога в теории случайных процессов. Расстояния, на которых значения анизотропного случайного поля становятся некоррелированными, различны по разным направлениям. Так, для анизотропной гауссовой корреляционной функции

$$\psi_{\zeta}(\mathbf{r}) = \sigma_{\zeta}^2 \exp\left[-\frac{1}{2}\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}\right)\right], \quad (2.29)$$

которая является частным случаем функций вида (2.28), величины  $a$ ,  $b$  и  $c$  характеризуют масштабы пространственной корреляции в направлениях  $x$ ,  $y$  и  $z$ . Масштаб  $l$  для функций вида (2.27) характеризует радиус корреляции в направлении, перпендикулярном плоскости  $\alpha x + \beta y + \gamma z = 0$ , тогда как в самой этой плоскости (и в параллельных ей) корреляция простирается до бесконечности.

Иногда вместо статистической анизотропии говорят об *анизомерии* случайных полей. Термин «анизомерные флуктуации» удобен в некоторых задачах электродинамики и теории упругости, в которых речь идет о статистически анизотропных флуктуациях параметров в анизотропных же средах, но пользоваться им нам почти не придется.

Наряду с моментами первого и второго порядков, можно ввести и высшие моменты комплексного случайного поля  $\xi$ . Принято определять  $(m+n)$ -точечный момент  $B_{m,n}$  (момент  $(m+n)$ -го порядка) следующим образом:

$$B_{m,n}(1, 2, \dots, m+n) = \langle \xi(1) \dots \xi(m) \xi^*(m+1) \dots \xi^*(m+n) \rangle. \quad (2.30)$$

Поле  $\xi$  входит под знак усреднения  $m$  раз, а комплексно сопряженное поле  $\xi^*$  —  $n$  раз. Для краткости вместо координат  $\mathbf{r}_j$  здесь снова написаны только номера точек  $j$ , например,  $\xi(m)$  вместо  $\xi(\mathbf{r}_m)$  и т. д. Моменты второго порядка  $B$  и  $\bar{B}$  запишутся в этих обозначениях следующим образом:

$$B(1, 2) = B_{1,1}(1, 2), \quad \bar{B}(1, 2) = B_{2,0}(1, 2). \quad (2.31)$$

Некоторые простые свойства моментов  $B_{m,n}$  вытекают непосредственно из определения. Так, очевидно, что  $B_{n,n}(1, \dots, n; n, \dots, 1) \geq 0$ . При перестановках внутри первой  $(1, \dots, m)$  и второй  $(m+1, \dots, m+n)$  групп аргументов момент  $B_{m,n}$  не меняется, а комплексное сопряжение приводит к следующим изменениям в порядке следования аргументов и индексов:

$$[B_{m,n}(1, \dots, m+n)]^* = B_{n,m}(m+n, \dots, 1). \quad (2.32)$$

Эрмитовость  $B$  и симметрия  $\bar{B}$  представляют собой частные случаи этого тождества.

### § 3. Пространственные спектральные разложения для однородных случайных полей

Запишем формальное разложение флуктуационной компоненты однородного случайного поля  $\tilde{\xi}(\mathbf{r})$  в трехкратный интеграл Фурье:

$$\tilde{\xi}(\mathbf{r}) = \int_{-\infty}^{\infty} \xi(\mathbf{k}) e^{i\mathbf{k}\mathbf{r}} d^3k. \quad (3.1)$$

Здесь  $\xi(\mathbf{k})$  — пространственная спектральная амплитуда (или, короче  $\mathbf{k}$ -амплитуда) поля  $\xi$ , которая выражается через  $\tilde{\xi}(\mathbf{r})$  при помощи обратного преобразования Фурье:

$$\xi(\mathbf{k}) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} \tilde{\xi}(\mathbf{r}) e^{-i\mathbf{k}\mathbf{r}} d^3\mathbf{r} \quad (3.2)$$

Иногда вместо  $\xi(\mathbf{k})$  используется более краткое обозначение  $\xi_{\mathbf{k}}$ . Спектральное разложение (3.1) мы назвали формальным потому, что для неубывающей на бесконечности функции, какой является однородное случайное поле  $\tilde{\xi}(\mathbf{r})$ , трансформанты Фурье



$\xi(\bar{\mathbf{k}})$  не существует. Для того чтобы придать спектральным разложениям таких случайных полей обычный математический смысл, следовало бы предположить, что поле однородно в большой, но конечной области  $V$ , вне которой оно достаточно быстро убывает до нуля, и переходить к бесконечным размерам области  $V$  лишь на последнем этапе расчетов, уже после статистического усреднения. Мы не будем, однако, пользоваться здесь такого рода приемами (как мы не делали этого в аналогичной ситуации и для спектральных разложений стационарных случайных процессов), а будем считать, что интегралы (3.1) и (3.2) существуют в смысле вероятностной сходимости, а именно—в среднем квадратичном (ч. I, § 40).

Согласно (3.1) функция корреляции однородного случайного поля  $\psi_{\xi}(\mathbf{r}_1 - \mathbf{r}_2)$  равна

$$\psi_{\xi}(\mathbf{r}_1 - \mathbf{r}_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle \xi(\mathbf{k}') \xi^*(\mathbf{k}'') \rangle e^{i(\mathbf{k}'\mathbf{r}_1 - \mathbf{k}''\mathbf{r}_2)} d^3\mathbf{k}' d^3\mathbf{k}''. \quad (3.3)$$

Спектральные амплитуды  $\xi(\mathbf{k}')$  и  $\xi^*(\mathbf{k}'')$  оказываются дельта-коррелированными по  $\mathbf{k}$ . В самом деле, в силу (3.2)

$$\langle \xi(\mathbf{k}') \xi^*(\mathbf{k}'') \rangle = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle \tilde{\xi}(\mathbf{r}_1) \tilde{\xi}^*(\mathbf{r}_2) \rangle e^{-i(\mathbf{k}'\mathbf{r}_1 - \mathbf{k}''\mathbf{r}_2)} d^3\mathbf{r}_1 d^3\mathbf{r}_2.$$

Введем здесь новые переменные интегрирования  $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ ,  $\mathbf{R} = (\mathbf{r}_1 + \mathbf{r}_2)/2$ , для которых  $d^3\mathbf{r}_1 d^3\mathbf{r}_2 = d^3\mathbf{r} d^3\mathbf{R}$ . Учитывая, что  $\langle \tilde{\xi}(\mathbf{r}_1) \tilde{\xi}^*(\mathbf{r}_2) \rangle = \psi_{\xi}(\mathbf{r}_1 - \mathbf{r}_2) = \psi_{\xi}(\mathbf{r})$  и  $\mathbf{k}'\mathbf{r}_1 - \mathbf{k}''\mathbf{r}_2 = (\mathbf{k}' - \mathbf{k}'')\mathbf{R} + (\mathbf{k}' + \mathbf{k}'')\mathbf{r}/2$ , после интегрирования по  $\mathbf{R}$  получаем

$$\langle \xi(\mathbf{k}') \xi^*(\mathbf{k}'') \rangle = \Phi_{\xi}(\mathbf{k}') \delta(\mathbf{k}' - \mathbf{k}''), \quad (3.4)$$

где

$$\Phi_{\xi}(\mathbf{k}) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} \psi_{\xi}(\mathbf{r}) e^{-i\mathbf{k}\mathbf{r}} d^3\mathbf{r}. \quad (3.5)$$

Подставляя же (3.4) в (3.3), находим

$$\psi_{\xi}(\mathbf{r}) = \int_{-\infty}^{\infty} \Phi_{\xi}(\mathbf{k}) e^{i\mathbf{k}\mathbf{r}} d^3\mathbf{k}. \quad (3.6)$$

Функция  $\Phi_{\xi}(\mathbf{k})$  называется *пространственной спектральной плотностью* (*k-плотностью*) или, короче, *пространственным спектром* случайного однородного поля  $\xi(\mathbf{r})$  и аналогична спектральной плотности ( $\omega$ -плотности)  $g(\omega)$  в теории стационарных случайных процессов. Соотношение (3.5) представляет собой обобщение теоремы Хинчина (ч. I, § 41) на случайные поля. Заметим, что наряду с трехмерными спектральными разложениями

иногда удобно использовать также одномерные и двумерные спектральные разложения (см. задачи 2—4 в конце главы). Нетрудно убедиться, что пространственный спектр  $\Phi_{\mathbf{r}}(\mathbf{k})$  является вещественной и неотрицательной величиной. Вещественность  $\Phi_{\mathbf{r}}(\mathbf{k})$  вытекает из (2.19), а неотрицательность — из положительной определенности корреляционной функции.

В отличие от  $\Phi_{\mathbf{r}}(\mathbf{k})$  взаимный пространственный спектр

$$\Phi_{\mu\nu}(\mathbf{k}) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} \psi_{\mu\nu}(\mathbf{r}) e^{-i\mathbf{k}\mathbf{r}} d^3r, \quad (3.7)$$

отвечающий взаимной корреляционной функции  $\psi_{\mu\nu}(\mathbf{r})$  двух комплексных полей  $\mu(\mathbf{r})$  и  $\nu(\mathbf{r})$ , в общем случае является комплексной величиной.

Для четных корреляционных функций  $\psi_{\mathbf{r}}(-\mathbf{r}) = \psi_{\mathbf{r}}(\mathbf{r})$ , в частности для функции корреляции вещественного или изотропного поля, спектральная плотность тоже четна,  $\Phi_{\mathbf{r}}(-\mathbf{k}) = \Phi_{\mathbf{r}}(\mathbf{k})$ . В этом случае  $\psi_{\mathbf{r}}(\mathbf{r})$  и  $\Phi_{\mathbf{r}}(\mathbf{k})$  связаны косинус-преобразованием Фурье:

$$\psi_{\mathbf{r}}(\mathbf{r}) = \int_{-\infty}^{\infty} \Phi_{\mathbf{r}}(\mathbf{k}) \cos \mathbf{k}\mathbf{r} d^3k, \quad (3.8)$$

$$\Phi_{\mathbf{r}}(\mathbf{k}) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} \psi_{\mathbf{r}}(\mathbf{r}) \cos \mathbf{k}\mathbf{r} d^3r. \quad (3.9)$$

В частном случае *изотропного* однородного случайного поля, для которого  $\psi_{\mathbf{r}}(\mathbf{r}) = \psi_{\mathbf{r}}(r)$ , пространственная спектральная плотность зависит только от модуля вектора  $\mathbf{k}$ . Для того чтобы это показать, перейдем в (3.5) к полярным координатам  $(r, \theta, \varphi)$  с полярной осью, направленной по вектору  $\mathbf{k}$ . Тогда  $\mathbf{k}\mathbf{r} = kr \cos \theta$  и мы получаем

$$\begin{aligned} \Phi_{\mathbf{r}}(\mathbf{k}) &= \frac{1}{(2\pi)^3} \int_0^{2\pi} d\varphi \int_0^{\pi} \sin \theta d\theta \int_0^{\infty} \psi_{\mathbf{r}}(r) e^{-i\mathbf{k}\mathbf{r} \cos \theta} r^2 dr = \\ &= \frac{1}{2\pi^2 k} \int_0^{\infty} \sin kr \psi_{\mathbf{r}}(r) r dr = \Phi_{\mathbf{r}}(k). \end{aligned} \quad (3.10)$$

Зависимость  $\Phi_{\mathbf{r}}$  только от  $k$  позволяет записать (3.6) и (3.8) в более простом виде, а именно через однократный интеграл по  $k$ . Вводя сферические координаты в пространстве волновых векторов  $\mathbf{k}$  (с полярной осью по вектору  $\mathbf{r}$ ), после интегрирования по угловым переменным получаем

$$\psi_{\mathbf{r}}(r) = \frac{4\pi}{r} \int_0^{\infty} \sin kr \Phi_{\mathbf{r}}(k) k dk. \quad (3.11)$$

Приведем в качестве примера значение пространственного спектра  $\Phi_{\xi}(k)$  однородного и изотропного случайного поля с гауссовой функцией корреляции (2.24):

$$\Phi_{\xi}(k) = (2\pi)^{-3/2} \sigma_{\xi}^2 l^3 e^{-1/2 k^2 l^2}, \quad (3.12)$$

и с экспоненциальной функцией корреляции (2.25):

$$\Phi_{\xi}(k) = \frac{\sigma_{\xi}^2 l^3}{\pi(1+k^2 l^2)^2}. \quad (3.13)$$

Эти выражения можно получить при помощи любой из формул (3.5), (3.9) или, что проще всего, (3.10).

Из этих примеров видно, что «неопределенность» (ширина) спектра  $\Delta k$  обратно пропорциональна радиусу корреляции случайного поля:  $\Delta k \sim 1/l$ . Но  $l$  характеризует «неопределенность» (ширину)  $\Delta r$  корреляционной функции, которая существенно отличается от нуля лишь при  $r \leq \Delta r \sim l$ . В результате для изотропных случайных полей можно формулировать соотношение *неопределенностей* (т. е. размытостей)

$$\Delta k \Delta r \geq 1, \quad (3.14)$$

которое является аналогом соотношения  $\Delta \omega \Delta t \geq 1$  для случайных процессов (ч. I, § 44). Согласно (3.14) коротко коррелированным полям ( $\Delta r \sim l$  мало) отвечают широкие пространственные спектры ( $\Delta k$  велико), тогда как при больших радиусах корреляции ширина спектра мала.

Для анизотропных полей неопределенности (радиусы корреляции) по разным направлениям неодинаковы, и для них вместо одного неравенства (3.14) выполняются сразу три неравенства:

$$\Delta k_x \Delta r_x \geq 1, \quad \Delta k_y \Delta r_y \geq 1, \quad \Delta k_z \Delta r_z \geq 1. \quad (3.15)$$

Для иллюстрации этих соотношений можно привести пространственный спектр, отвечающий гауссовой корреляционной функции (2.29):

$$\Phi_{\xi}(\mathbf{k}) = \frac{\sigma_{\xi}^2 abc}{(2\pi)^{3/2}} \exp \left\{ -\frac{1}{2} (a^2 k_x^2 + b^2 k_y^2 + c^2 k_z^2) \right\}. \quad (3.16)$$

Ширина этого спектра по осям  $k_x$ ,  $k_y$ ,  $k_z$  обратно пропорциональна соответственно величинам  $a$ ,  $b$  и  $c$ , которые характеризуют масштаб изменения корреляционной функции (2.29) по осям  $x$ ,  $y$  и  $z$ .

Для существования  $\Psi_{\xi}(\mathbf{r})$ , т. е. для существования интеграла в правой части (3.8) и (3.11), необходимо, с одной стороны, чтобы с ростом  $k$  спектральная плотность  $\Phi_{\xi}(\mathbf{k})$  убывала быстрее  $k^3$ . С другой стороны, при  $k \rightarrow 0$  у спектра  $\Phi_{\xi}(\mathbf{k})$  допустимо наличие степенной особенности вида  $k^{-\alpha}$  с  $\alpha < 3$ .

В тесной связи с условиями существования корреляционной функции находится вопрос о требованиях, которым должен подчиняться спектр *дифференцируемых* случайных полей. Так же, как и в теории случайных процессов (ч. I, § 19), можно убедиться, что необходимое и достаточное условие существования (разумеется, в среднем квадратичном) первых пространственных производных, например  $\nabla \zeta(\mathbf{r})$ , сводится к существованию величины

$$\langle \nabla \zeta(\mathbf{r}_1) \nabla \zeta^*(\mathbf{r}_2) \rangle = -\Delta \psi_{\zeta}(\mathbf{r}_1 - \mathbf{r}_2). \quad (3.17)$$

Если подставить сюда спектральное разложение (3.6) и положить  $\mathbf{r}_1 = \mathbf{r}_2 = \mathbf{r}$ , то условие существования  $\nabla \zeta$  запишется в виде

$$\langle |\nabla \zeta(\mathbf{r})|^2 \rangle = -\Delta \psi_{\zeta}(0) = \int_{-\infty}^{\infty} k^2 \Phi_{\zeta}(\mathbf{k}) d^3k < \infty. \quad (3.18)$$

Это условие допускает при  $k \rightarrow 0$  степенные особенности вида  $\Phi_{\zeta}(\mathbf{k}) \sim k^{-\alpha}$  более высокого порядка, чем это требуется для существования  $\psi_{\zeta}(\mathbf{r})$ : интеграл (3.18) сходится, если  $\alpha < 5$ , тогда как для сходимости интеграла (3.6) допустимы лишь  $\alpha < 3$ . В то же время условие (3.18) предъявляет более жесткое требование к скорости убывания  $\Phi_{\zeta}(\mathbf{k})$  при  $k \rightarrow \infty$ . Необходимо, чтобы  $\Phi_{\zeta}(\mathbf{k})$  с ростом  $k$  убывало, как  $k^{-\mu}$  с  $\mu > 5$  (для существования  $\psi_{\zeta}(\mathbf{r})$  необходимо лишь  $\mu > 3$ ). Последнему условию не удовлетворяет, например, случайное поле с экспоненциальной функцией корреляции (2.25), поскольку его спектральная плотность (3.13) убывает на бесконечности недостаточно быстро (как  $k^{-4}$ ). Следовательно, такое поле недифференцируемо.

Еще более жесткие требования при  $k \rightarrow \infty$  предъявляет к  $\Phi_{\zeta}(\mathbf{k})$  существование у случайного поля производных  $n$ -го порядка. Из условия  $\langle |\nabla^n \zeta(\mathbf{r})|^2 \rangle < \infty$  следует, что должно выполняться неравенство

$$\int k^{2n} \Phi_{\zeta}(\mathbf{k}) d^3k < \infty, \quad (3.19)$$

так что с ростом  $k$  спектр  $\Phi_{\zeta}(\mathbf{k})$  должен убывать быстрее, чем  $k^{-3-2n}$ . Очевидно, для бесконечно дифференцируемого поля  $\Phi_{\zeta}(\mathbf{k})$  должно уменьшаться при  $k \rightarrow \infty$  быстрее любой отрицательной степени  $k$ , например, экспоненциально. Таким является, например, случайное поле с гауссовой корреляционной функцией (2.24).

Примером спектра, спадающего более медленно (чем по экспоненте), но быстрее любого  $1/k^m$ , может служить функция

$$\Phi_{\zeta}(\mathbf{k}) = C \exp \{ -(\ln k)^2 \} = C k^{-\ln k},$$

для которой при любом конечном  $m$

$$\int_0^{\infty} k^m \Phi_{\zeta}(k) dk = \sqrt{\pi} \exp \left\{ \frac{(m+1)^2}{4} \right\} < \infty.$$

#### § 4. Локально однородные случайные поля

Случайное комплексное поле  $\zeta(\mathbf{r})$  можно характеризовать не только корреляционной функцией (2.1), но и так называемой *структурной функцией*, которая представляет собой средний квадрат модуля *приращения* флуктуационной компоненты  $\tilde{\zeta}(\mathbf{r})$  поля  $\zeta(\mathbf{r})$ :

$$D_{\zeta}(\mathbf{r}_1, \mathbf{r}_2) = \langle |\tilde{\zeta}(\mathbf{r}_1) - \tilde{\zeta}(\mathbf{r}_2)|^2 \rangle. \quad (4.1)$$

Очевидно, при  $\mathbf{r}_1 = \mathbf{r}_2$  структурная функция обращается в нуль,  $D_{\zeta}(\mathbf{r}_1, \mathbf{r}_1) = 0$ .

Если структурная функция  $D_{\zeta}(\mathbf{r}_1, \mathbf{r}_2)$  и приращение среднего поля  $f_{\zeta}(\mathbf{r}_1, \mathbf{r}_2) = \langle \zeta(\mathbf{r}_1) \rangle - \langle \zeta(\mathbf{r}_2) \rangle$  зависят только от разности  $\mathbf{r}_1 - \mathbf{r}_2$ :

$$D_{\zeta}(\mathbf{r}_1, \mathbf{r}_2) = D_{\zeta}(\mathbf{r}_1 - \mathbf{r}_2), \quad f_{\zeta}(\mathbf{r}_1, \mathbf{r}_2) = f_{\zeta}(\mathbf{r}_1 - \mathbf{r}_2), \quad (4.2)$$

то такие поля называют *локально однородными*. Понятие локально однородных полей ввел А. Н. Колмогоров [3, 4], а термин «структурная функция» был предложен позднее А. М. Обуховым. Условия локальной однородности налагают определенные ограничения на моменты *приращений*  $\tilde{\zeta}(\mathbf{r})$ , т. е. *разностей* значений поля в двух точках  $\mathbf{r}_1$  и  $\mathbf{r}_2$ , а не на моменты самих этих значений  $\zeta(\mathbf{r}_1)$  и  $\zeta(\mathbf{r}_2)$ . В силу полной аналогии таких полей со случайными процессами со стационарными (первыми) приращениями (ч. I, § 5б) можно было бы назвать локально однородное поле также *случайным полем с однородными приращениями*, но это название не получило распространения.

Требованию однородности величины  $f_{\zeta}(\mathbf{r}_1, \mathbf{r}_2) = \langle \zeta(\mathbf{r}_1) \rangle - \langle \zeta(\mathbf{r}_2) \rangle$  можно удовлетворить, только если среднее поле линейно зависит от  $\mathbf{r}$ :  $\langle \zeta(\mathbf{r}) \rangle = \langle \mathbf{a} \rangle \mathbf{r} + \text{const}$ , где  $\mathbf{a}$  — произвольный вектор, который может быть и не случайным. Действительно, при линейной зависимости  $\langle \zeta(\mathbf{r}) \rangle$  от  $\mathbf{r}$  имеем

$$f_{\zeta}(\mathbf{r}_1, \mathbf{r}_2) = \langle \mathbf{a} \rangle (\mathbf{r}_1 - \mathbf{r}_2) = f_{\zeta}(\mathbf{r}_1 - \mathbf{r}_2).$$

Если у поля существует корреляционная функция  $\psi_{\zeta}(\mathbf{r}_1, \mathbf{r}_2) = \langle \tilde{\zeta}(\mathbf{r}_1) \tilde{\zeta}^*(\mathbf{r}_2) \rangle$ , то, согласно (4.1),

$$D_{\zeta}(\mathbf{r}_1, \mathbf{r}_2) = \psi_{\zeta}(\mathbf{r}_1, \mathbf{r}_1) + \psi_{\zeta}(\mathbf{r}_2, \mathbf{r}_2) - \psi_{\zeta}(\mathbf{r}_1, \mathbf{r}_2) - \psi_{\zeta}^*(\mathbf{r}_1, \mathbf{r}_2). \quad (4.3)$$

В частности, для однородного поля  $\zeta(\mathbf{r})$  имеем из (4.3)

$$D_{\zeta}(\mathbf{r}) = 2[\psi_{\zeta}(\mathbf{0}) - \psi_{\zeta}(\mathbf{r})]. \quad (4.4)$$

Важным преимуществом структурной функции является то, что она может сохранять смысл и в том более общем случае, когда корреляционной функции не существует. Ситуация здесь такая же, как для случайных процессов (стационарные процессы пред-

ставляют собой частный случай процессов со стационарными приращениями). Аналогия становится еще более полной, если речь идет о *локально однородных и изотропных полях*. Под этим понимаются случайные поля, у которых  $D_{\xi}(r)$  и  $f_{\xi}(r)$  зависят только от модуля вектора  $r = r_1 - r_2$ :

$$D_{\xi}(r) = D_{\xi}(r), \quad f_{\xi}(r) = f_{\xi}(r). \quad (4.5)$$

Второе из этих условий удовлетворяется, если  $\langle a \rangle = 0$ , т. е. вектор  $a$  обладает изотропным распределением в пространстве. В приложениях и, в частности, в теории атмосферной турбулентности часто можно считать равным нулю сам вектор  $a$ . Так или иначе, среднее значение локально однородного и изотропного поля постоянно:  $\langle \xi \rangle = \text{const}$ . Для однородного и изотропного случайного поля равенство (4.4) принимает вид

$$D_{\xi}(r) = 2 [\psi_{\xi}(0) - \psi_{\xi}(r)], \quad (4.6)$$

причем  $D_{\xi}(r)$  и  $\psi_{\xi}(r)$  вещественны. Если  $\psi_{\xi}(r)$  при  $r \rightarrow \infty$  исчезает, то  $D_{\xi}(\infty) = 2\psi_{\xi}(0)$ , что позволяет выразить  $\psi_{\xi}(r)$  через  $D_{\xi}(r)$ :

$$\psi_{\xi}(r) = 1/2 [D_{\xi}(\infty) - D_{\xi}(r)]. \quad (4.7)$$

Важное свойство структурной функции состоит в том, что на нее не влияют большие по пространственной протяженности флуктуации  $\xi$ , т. е. флуктуации с характерным размером  $l \gg r = |r_1 - r_2|$ . Обусловленные такими флуктуациями возмущения практически одинаковы в точках  $r_1$  и  $r_2$ , разность  $\xi(r_1) - \xi(r_2)$  для них мала и, соответственно, мал их вклад в  $D_{\xi}(r)$ . Корреляционная же функция в равной мере учитывает флуктуации любого масштаба. Именно поэтому использование структурной, а не корреляционной функции оказывается физически оправданным в тех случаях, когда крупномасштабные флуктуации поля не сказываются на интересующих нас явлениях. Это вовсе не означает, что такие флуктуации отсутствуют. Напротив, их доля в результирующих флуктуационных возмущениях может быть даже велика, но для наблюдаемых явлений они несут незначительный вклад.

Примером может служить статистическая теория развитой турбулентности [1, 2], т. е. такого вихревого движения газа или жидкости, в котором присутствуют вихри с очень широким диапазоном размеров  $l$ . Наиболее интересными по своим внутренним закономерностям здесь оказываются вихри, значительно уступающие по размерам тем наиболее крупным вихрям (размера  $L_0$ , так называемого *внешнего масштаба турбулентности*), которые порождены первичным потоком и еще сильно зависят от его геометрических и кинематических особенностей. Именно для субдиапазона  $l \ll L_0$  А. Н. Колмогоров ввел понятие локальной однородности случайного поля и предложил для его статисти-

ческого описания функцию  $D_{\epsilon}$ , которая просто исключает крупномасштабные неоднородности из рассмотрения.

Тем самым, предположение об однородности, если оно делается для  $D_{\epsilon}$ , гораздо менее обременительно, т. е. оно позволяет охватить класс случайных полей (локально однородные поля) более широкий, чем такое же предположение для  $\Psi_{\epsilon}$  (однородные поля).

Обратимся к пространственным спектральным разложениям локально однородных полей. По аналогии со спектральными разложениями для процессов со стационарными случайнымиращениями (ч. I, § 56) имеем

$$D_{\epsilon}(r) = 2 \int_{-\infty}^{\infty} \Phi_{\epsilon}(k) (1 - \cos kr) d^3k. \quad (4.8)$$

Интеграл (4.8) сходится, если при  $k \rightarrow 0$  спектр  $\Phi_{\epsilon}(k)$  имеет степенную особенность вида  $k^{-\alpha}$  с  $\alpha < 5$ , тогда как существование  $\Psi_{\epsilon}(r)$  обеспечено лишь при  $\alpha < 3$ .

Формула обращения разложения (4.8) в случае локально однородных полей получается несколько более сложным путем, чем для однородных полей. Сначала надо продифференцировать (4.8) по  $r$  и только после этого применить обратное преобразование Фурье. Мы получаем тогда, что

$$\Phi_{\epsilon}(k) = \frac{1}{16\pi^3 k^3} \int_{-\infty}^{\infty} k \nabla D_{\epsilon}(r) \sin kr d^3r. \quad (4.9)$$

Для изотропного поля, переходя в (4.8) и (4.9) к сферическим координатам и учитывая, что  $k \nabla D_{\epsilon}(r) = kr D'_{\epsilon}(r)/r$  (штрих обозначает производную по  $r$ ), находим

$$D_{\epsilon}(r) = 8\pi \int_0^{\infty} \left(1 - \frac{\sin kr}{kr}\right) \Phi_{\epsilon}(k) k^2 dk, \quad (4.10)$$

$$\Phi_{\epsilon}(k) = \frac{1}{4\pi^3 k^3} \int_0^{\infty} (\sin kr - kr \cos kr) D'_{\epsilon}(r) dr. \quad (4.11)$$

Рассмотрим в качестве примера локально однородного и изотропного поля пространственные флуктуации диэлектрической проницаемости в турбулентной атмосфере. Структурная функция этих флуктуаций

$$D_{\epsilon}(r) = \langle |\bar{\epsilon}(r_1) - \bar{\epsilon}(r_2)|^2 \rangle$$

подчиняется при достаточно больших  $r$  «закону двух третей» Колмогорова—Обухова [1]:

$$D_{\epsilon}(r) \approx C_{\epsilon}^2 r^{2/3}, \quad r \gg l_0, \quad (4.12)$$

где  $C_\varepsilon$  — так называемая *структурная постоянная*, а  $l_0$  — *внутренний масштаб турбулентности*<sup>1)</sup>. При малых же  $r$ , т. е.  $r \ll l_0$ , структурная функция нарастает по квадратичному закону:

$$D_\varepsilon(r) \approx C_\varepsilon^2 l_0^{-4/3} r^2, \quad r \ll l_0. \quad (4.13)$$

Внутренний масштаб  $l_0$  входит в формулы (4.12) и (4.13) так, что значения обоих асимптотических выражений одинаковы при

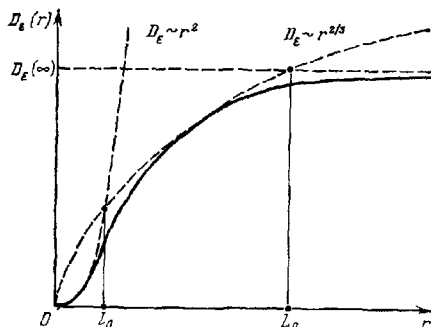


Рис. 1.

$r = l_0$ . Соответствующие графики показаны на рис. 1 пунктирными кривыми, а сплошной линией изображен реальный ход структурной функции  $D_\varepsilon(r)$ .

Подобрать спектр, отвечающий реальному ходу структурной функции  $D_\varepsilon(r)$ , можно из следующих соображений. Для степенной структурной функции

$$D(r) = C^2 r^\mu, \quad 0 < \mu < 2, \quad (4.14)$$

пространственная спектральная плотность равна [1]

$$\Phi(k) = \frac{\Gamma(\mu+2)}{4\pi^2} \sin\left(\frac{\pi\mu}{2}\right) C^2 k^{-(\mu+3)}, \quad (4.15)$$

так что «закону двух третей» ( $\mu = 2/3$ ) отвечает спектр

$$\Phi_\varepsilon(k) = \frac{\Gamma(4/3) \sin \pi/3}{4\pi^2} C_\varepsilon^2 k^{-11/3} = 0,033 C_\varepsilon^2 k^{-11/3}. \quad (4.16)$$

Для того чтобы получить квадратичный ход  $D_\varepsilon(r)$  при малых значениях  $r \ll l_0$ , следует подавить спектральную плотность (4.16)

<sup>1)</sup> В противоположность внешнему масштабу  $L_0$  внутренний масштаб  $l_0$  — это *наименьший* размер вихрей. Возмущения с  $l < l_0$  уже ламинарны.



при больших значениях  $k \gg 1/l_0$ . Физически эта операция отображает диссипацию турбулентных вихрей из-за вязкости, когда их масштаб  $1/k$  становится малым ( $1/k \ll l_0$ ). Подавление плотности (4.16) при больших  $k$  можно осуществить, например, путем введения множителя  $e^{-k^2/\kappa_m^2}$ , т. е. полагая

$$\Phi_e(k) = 0,033 C_e^2 k^{-11/2} e^{-k^2/\kappa_m^2}. \quad (4.17)$$

Получаемая отсюда структурная функция  $D_e(r)$  будет обладать требуемой асимптотикой (4.12), (4.13), если параметр обрезания  $\kappa_m$  взять равным  $5,92/l_0$ . Она хорошо описывает турбулентное поле при не слишком больших значениях  $r$  — не превышающих внешнего масштаба турбулентности  $L_0$ . В действительности же при  $r \gg L_0$  структурная функция «насыщается» (см. рис. 1) и стремится к конечному значению, которое удобно записать в виде

$$D_e(\infty) = C_e^2 L_0^{2/3}. \quad (4.18)$$

Если значение  $L_0$  вводится таким путем, то кривая (4.12) пересекает уровень  $D_e = D_e(\infty)$  как раз при  $r = L_0$ . Согласно (4.6) предельное значение (4.18) равно  $2\sigma_e^2$  — удвоенному среднему квадрату флуктуаций:  $C_e^2 L_0^{2/3} = 2\sigma_e^2$ .

В результате мы приходим к тому, что структурную функцию  $D_e(r)$  можно аппроксимировать на отдельных интервалах  $r$  следующими функциями:

$$D(r) = \begin{cases} C_e^2 l_0^{-4/3} r^2 = 2\sigma_e^2 r^2 l_0^{-4/3} L_0^{-2/3}, & r \ll l_0, \\ C_e^2 r^{2/3} = 2\sigma_e^2 (r/L_0)^{2/3}, & l_0 \ll r \ll L_0, \\ C_e^2 L_0^{2/3} = 2\sigma_e^2, & L_0 \ll r. \end{cases} \quad (4.19)$$

Ограниченность структурной функции при  $r \rightarrow \infty$  можно отразить и в (4.17), заменив множитель  $k^{-11/2}$  на  $(k^2 + \kappa_0^2)^{-11/2}$ , где  $\kappa_0 = 2\pi/L_0$  — волновое число, отвечающее внешнему масштабу турбулентности. При такой замене, т. е. при спектральной плотности

$$\Phi_e(k) = 0,033 C_e^2 (k^2 + \kappa_0^2)^{-11/2} e^{-k^2/\kappa_m^2}, \quad (4.20)$$

значения интегралов (4.8) и (4.10) будут конечны при любых  $r$ .

Конечно, поведение флуктуаций  $\epsilon$  в реальной турбулентной атмосфере подчинено более сложным закономерностям, чем приведенные здесь аппроксимации. Тем не менее формула (4.20) достаточно хорошо описывает пространственный спектр турбулентных флуктуаций  $\epsilon$  во многих задачах радиофизики.

Систематическое изложение вопросов, относящихся к статистической гидродинамике, можно найти в монографиях [1, 2].

## § 5. Квазиоднородные поля

Идеализация реальных неоднородных случайных полей при помощи рассмотренных выше локально однородных полей, описывающих довольно специальный вид уклонений от однородности, далеко не всегда применима. Поэтому особое значение приобретает рассмотрение и других специальных случаев нарушения однородности. Сюда относится, в частности, класс полей, неоднородность которых в определенном смысле пространственно *медленна*, или, иначе говоря, *плавна*, а в остальном произвольна. Это так называемые статистически *квазиоднородные* поля, т. е. поля с плавно меняющимися статистическими характеристиками — средними значениями, дисперсиями и т. д. Очевидно, они представляют собой пространственный аналог квазистационарных случайных процессов, рассмотренных в ч. I, § 57.

Согласно (2.6) в общем случае корреляционная функция связана с коэффициентом корреляции  $K_{\xi}$  соотношением

$$\psi_{\xi}(\mathbf{r}_1, \mathbf{r}_2) = \sigma_{\xi}^2(\mathbf{r}_1) \sigma_{\xi}^2(\mathbf{r}_2) K_{\xi}(\mathbf{r}_1, \mathbf{r}_2),$$

которое в координатах  $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$  и  $\mathbf{R} = (\mathbf{r}_1 + \mathbf{r}_2)/2$  принимает вид

$$\psi_{\xi}(\mathbf{r}_1, \mathbf{r}_2) = \sigma_{\xi}^2(\mathbf{R} + \mathbf{r}/2) \sigma_{\xi}^2(\mathbf{R} - \mathbf{r}/2) K_{\xi}(\mathbf{r}, \mathbf{R}). \quad (5.1)$$

Если  $l_{\xi}$  — радиус корреляции поля  $\xi$  (масштаб изменения  $K_{\xi}(\mathbf{r}, \mathbf{R})$  по разностному аргументу  $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ ), а  $L$  — характерный масштаб изменения дисперсии  $\sigma_{\xi}^2$ , среднего поля  $\langle \xi \rangle$  и коэффициента корреляции  $K_{\xi}$  по аргументу  $\mathbf{R}$ , то для квазиоднородных полей выполняется неравенство

$$l_{\xi} \ll L. \quad (5.2)$$

В силу (5.2) можно заменить в формуле (5.1) произведение  $\sigma_{\xi}^2(\mathbf{R} + \mathbf{r}/2) \sigma_{\xi}^2(\mathbf{R} - \mathbf{r}/2)$  на  $\sigma_{\xi}^2$ . Такая замена безусловно справедлива внутри сферы радиуса  $r \leq l_{\xi}$ ; что же касается пространства вне этой сферы, то там значение множителя при  $K_{\xi}$  несущественно, так как  $K_{\xi}(\mathbf{r}, \mathbf{R}) \rightarrow 0$  при  $r \gg l_{\xi}$ . Таким образом, мы приходим к следующему выражению для функции корреляции квазиоднородного поля:

$$\psi_{\xi}(\mathbf{r}, \mathbf{R}) = \sigma_{\xi}^2(\mathbf{R}) K_{\xi}(\mathbf{r}, \mathbf{R}). \quad (5.3)$$

Поскольку квазиоднородность представляет собой нарушение строгой однородности, дельта-корреляция спектральных амплитуд  $\xi(\mathbf{k})$  теперь уже не имеет места, подобно тому как нет дельта-корреляции для  $\omega$ -амплитуд в случае квазистационарных случайных процессов (ч. I, § 57).

Рассматривая спектральные разложения таких процессов (ч. I, § 57), мы выяснили, однако, что в случае квазистационарных процессов определенный физический смысл приобретает *мгновен-*

ный частотный спектр  $g(\omega, t)$ , получающийся при спектральном разложении функции корреляции  $\psi(\tau, t)$  только по «быстрой» переменной  $\tau = t_1 - t_2$ . То же справедливо и для квазиоднородных полей. Записав преобразование Фурье только по разностной переменной  $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ :

$$\psi_{\mathbf{r}}(\mathbf{r}, \mathbf{R}) = \int_{-\infty}^{\infty} \Phi_{\mathbf{r}}(\mathbf{k}, \mathbf{R}) e^{i\mathbf{k}\mathbf{r}} d^3k, \quad (5.4)$$

можно рассматривать величину

$$\Phi_{\mathbf{r}}(\mathbf{k}, \mathbf{R}) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} \psi_{\mathbf{r}}(\mathbf{r}, \mathbf{R}) e^{-i\mathbf{k}\mathbf{r}} d^3r \quad (5.5)$$

как *локальную* пространственную спектральную плотность, которая медленно меняется от одной точки  $\mathbf{R}$  к другой (с характерным масштабом  $L \gg l_c$ ).

Естественно, напрашивается сочтание обоих видов нарушения пространственной однородности, т. е. локальной однородности и квазиоднородности. Такого рода поля можно было бы назвать *локально квазиоднородными*, определив их как поля, *первые приращения которых квазиоднородны*. Примером локально квазиоднородного (и изотропного) поля может служить рассмотренное в [1] поле флуктуаций диэлектрической проницаемости  $\tilde{\epsilon}(\mathbf{r})$  в плавно неоднородной по высоте  $z$  турбулентной атмосфере.

## § 6. Пространственно-временные спектральные разложения случайных полей

До сих пор мы отвлекались от временной зависимости случайных полей с тем, чтобы оттенить особенности, обусловленные зависимостью полей от нескольких пространственных координат и отсутствующие у однопараметрических случайных функций.

Конечно, при учете зависимости случайных полей также от времени возникает ряд новых возможностей в отношении различных комбинаций пространственных и временных статистических свойств поля  $\xi(t, \mathbf{r})$  (стационарности по  $t$ , однородности по  $\mathbf{r}$  и разных видов отклонений от стационарности и однородности), о чем мы уже упоминали в § 1. Однако использование четырехмерных (пространственно-временных) гармонических разложений  $\xi(t, \mathbf{r})$  с формальной стороны осуществляется довольно очевидным образом и соответствующие обобщения вряд ли требуют подробного описания. Мы приведем лишь некоторые наиболее важные соотношения и формулы, которые понадобятся нам в дальнейшем.

Запишем для флуктуаций одномерного случайного поля  $\tilde{\xi}(t, r)$  четырехмерное спектральное разложение:

$$\tilde{\xi}(t, r) = \iint \xi(\omega, k) e^{i(kr - \omega t)} d\omega d^3k, \quad (6.1)$$

где  $\xi(\omega, k)$  —  $\omega k$ -амплитуда (точнее, амплитудная плотность) поля, иногда обозначаемая  $\xi_{\omega k}$ . В отличие от ч. I, где в гармонических разложениях процессов вводился временной фактор  $e^{i\omega t}$ , в данной части книги зависимость от времени мы будем описывать множителем  $e^{-i\omega t}$ , так как при рассмотрении волновых явлений удобнее оперировать с пространственным гармоническим множителем  $e^{ikr}$ , а не  $e^{-ikr}$ . В связи с этим при использовании тех или иных комплексных выражений из ч. I надо заменять в них  $i\omega$  на  $-i\omega$ .

У однородного и стационарного случайного поля спектральные амплитуды  $\xi(\omega', k')$  и  $\xi(\omega'', k'')$  дельта-коррелированы как по  $\omega$ , так и по  $k$ :

$$\langle \xi(\omega', k') \xi^*(\omega'', k'') \rangle = G_{\xi}(\omega', k') \delta(\omega' - \omega'') \delta(k' - k''). \quad (6.2)$$

Здесь  $G_{\xi}(\omega, k)$  — пространственно-временной спектр поля (или  $\omega k$ -плотность), который всегда неотрицателен. Иногда представляют этот спектр в виде  $G_{\xi}(\omega, k) = |\xi_{\omega k}|^2$ , но следует помнить, что это лишь иное обозначение множителя при дельта-функциях.

Пространственно-временная корреляционная функция  $\Psi_{\xi}(\tau, r)$  связана с  $\omega k$ -плотностью  $G_{\xi}(\omega, k)$  преобразованием Фурье:

$$\begin{aligned} \Psi_{\xi}(\tau, r) &\equiv \langle \tilde{\xi}(t_1, r_1) \tilde{\xi}^*(t_1 - \tau, r_1 - r) \rangle = \\ &= \iint_{-\infty}^{\infty} G_{\xi}(\omega, k) e^{i(kr - \omega\tau)} d\omega d^3k \end{aligned} \quad (6.3)$$

(обобщение теоремы Хинчина). Обратное преобразование Фурье:

$$G_{\xi}(\omega, k) = \frac{1}{(2\pi)^4} \iint_{-\infty}^{\infty} \Psi_{\xi}(\tau, r) e^{-i(kr - \omega\tau)} d\tau d^3r. \quad (6.4)$$

В случае четной корреляционной функции,  $\Psi(-\tau, -r) = \Psi(\tau, r)$ , экспоненту в этих разложениях можно заменить на  $\cos(kr - \omega\tau)$ .

Через  $\omega k$ -плотность  $G_{\xi}(\omega, k)$  можно выразить как пространственный  $\Phi_{\xi}(k)$ , так и временной (частотный)  $g_{\xi}(\omega)$  спектры случайного поля:

$$\Phi_{\xi}(k) = \int_{-\infty}^{\infty} G_{\xi}(\omega, k) d\omega = \int_{-\infty}^{\infty} \Psi_{\xi}(0, r) e^{-ikr} d^3r, \quad (6.5)$$

$$g_{\xi}(\omega) = \int_{-\infty}^{\infty} G_{\xi}(\omega, k) d^3k = \int_{-\infty}^{\infty} \Psi_{\xi}(\tau, 0) e^{i\omega\tau} d\tau. \quad (6.6)$$

В качестве примера рассмотрим поле с так называемыми «замороженными» неоднородностями, т. е. поле, в котором все временные изменения  $\xi(t, \mathbf{r})$  обусловлены только перемещением пространственных возмущений  $v(\mathbf{r})$  с постоянной скоростью  $\mathbf{v}$ . Если покоящееся однородное случайное поле  $v(\mathbf{r})$  описывается пространственной корреляционной функцией  $\Psi_v(\mathbf{r})$ , то при нерелятивистских значениях скорости <sup>1)</sup>  $\xi(t, \mathbf{r}) = v(\mathbf{r} - \mathbf{v}t)$ , и тогда

$$\Psi_{\xi}(\tau, \mathbf{r}) = \Psi_v(\mathbf{r} - \mathbf{v}\tau). \quad (6.7)$$

Подставляя (6.7) в (6.4), находим простую связь между  $\omega \mathbf{k}$ -плотностью  $G_{\xi}(\omega, \mathbf{k})$  «замороженного» поля  $\xi(t, \mathbf{r})$  и пространственной спектральной плотностью  $\Phi_v(\mathbf{k})$  покоящегося поля  $v(\mathbf{r})$ :

$$G_{\xi}(\omega, \mathbf{k}) = \Phi_v(\mathbf{k}) \delta(\omega - \mathbf{k}\mathbf{v}). \quad (6.8)$$

При помощи (6.5) и (6.8) легко убедиться, что

$$\Phi_{\xi}(\mathbf{k}) = \Phi_v(\mathbf{k}), \quad (6.9)$$

т. е. пространственные спектры полей  $\xi(t, \mathbf{r})$  и  $v(\mathbf{r})$  совпадают. Что же касается временного спектра флуктуаций  $\xi(t, \mathbf{r})$ , то из (6.6) и (6.8) находим

$$\begin{aligned} g_{\xi}(\omega) &= \int_{-\infty}^{\infty} dk_{\parallel} \int_{-\infty}^{\infty} d^2 k_{\perp} \Phi_v(k_{\parallel}, \mathbf{k}_{\perp}) \delta(\omega - k_{\parallel}v) = \\ &= \frac{1}{v} \int_{-\infty}^{\infty} \Phi_v\left(\frac{\omega}{v}, \mathbf{k}_{\perp}\right) d^2 k_{\perp}, \end{aligned} \quad (6.10)$$

где  $k_{\parallel}$  и  $\mathbf{k}_{\perp}$  — продольная и поперечная (по отношению к вектору скорости  $\mathbf{v}$ ) составляющие волнового вектора  $\mathbf{k}$ . Выражение (6.10) упрощается в случае изотропного поля  $v(\mathbf{r})$ , когда  $\Phi_v(k_{\parallel}, \mathbf{k}_{\perp}) = \Phi_v(V \sqrt{k_{\parallel}^2 + k_{\perp}^2})$ . Переходя в (6.10) к полярным координатам и вводя вместо  $k_{\perp}$  новую переменную интегрирования  $\kappa = V(\omega/v)^2 + k_{\perp}^2$ , получаем

$$g_{\xi}(\omega) = \frac{2\pi}{v} \int_{|\omega|/v}^{\infty} \Phi_v(\kappa) \kappa d\kappa. \quad (6.11)$$

Дифференцируя (6.11) по  $\omega$ , нетрудно выразить пространственный спектр замороженного изотропного поля  $\Phi_{\xi}(k) = \Phi_v(k)$  через частотный спектр  $g_{\xi}(\omega)$ :

$$\Phi_{\xi}(k) = - \frac{v^2}{2\pi k} \frac{dg_{\xi}(\omega)}{d\omega} \Big|_{\omega = kv}. \quad (6.12)$$

<sup>1)</sup> Об особенностях релятивистского случая см. в задаче 7.

Формулы типа (6.11) и (6.12) часто используются при анализе полей, которые можно с достаточной точностью считать замороженными.

Сказанное выше относительно спектральных разложений одномерных полей легко переносится на многомерное поле  $\xi^{(i)}(t, \mathbf{r})$  ( $i = 1, 2, \dots, n$ ). Наиболее простыми свойствами обладают поля *однородные* и *однородно связанные* (в широком смысле) в пространстве  $Q = (t, \mathbf{r})$ .

В случае *векторного* поля  $\xi^{(i)}$  элементы корреляционной матрицы  $\Psi_{ij}^{(\xi)}$  трансформируются при ортогональных преобразованиях координат как компоненты  $n$ -мерного тензора второго ранга, т. е. можно говорить о *корреляционном тензоре* векторного поля. Для однородных и изотропных векторных полей доказан ряд изящных теорем, часть которых сформулирована в задачах 8—10.

Разложение многомерного случайного поля  $\xi^{(i)}(t, \mathbf{r})$  в интеграл Фурье имеет вид, аналогичный (6.1); при этом для однородного и стационарного поля выполняется равенство

$$\langle \xi_i(\omega', \mathbf{k}') \xi_j^*(\omega'', \mathbf{k}'') \rangle = G_{ij}^{(\xi)}(\omega', \mathbf{k}') \delta(\omega' - \omega'') \delta(\mathbf{k}' - \mathbf{k}''), \quad (6.13)$$

которое естественно приводит к обобщению теоремы Хинчина на многомерные пространственно-временные поля:

$$\Psi_{ij}^{(\xi)}(\tau, \mathbf{r}) = \iint_{-\infty}^{\infty} G_{ij}^{(\xi)}(\omega, \mathbf{k}) e^{i(\mathbf{k}\mathbf{r} - \omega\tau)} d\omega d^3k \quad (6.14)$$

с обратным преобразованием

$$G_{ij}^{(\xi)}(\omega, \mathbf{k}) = \frac{1}{(2\pi)^4} \iint_{-\infty}^{\infty} \Psi_{ij}^{(\xi)}(\tau, \mathbf{r}) e^{-i(\mathbf{k}\mathbf{r} - \omega\tau)} d\tau d^3r. \quad (6.15)$$

Корреляционная матрица многомерного поля  $\Psi_{ij}^{(\xi)}$  положительно определена. Отсюда вытекает, что диагональные компоненты матрицы  $\omega\mathbf{k}$ -плотности вещественны и неотрицательны:  $G_{ii}^{(\xi)}(\omega, \mathbf{k}) \geq 0$ . Недиagonальные же элементы в общем случае комплексны.

Отметим теперь следующее существенное обстоятельство. Область интегрирования по  $\omega$  и  $\mathbf{k}$  в разложениях Фурье самих полей, равно как и их корреляционных функций (6.3) и (6.14), вообще говоря, четырехмерна. Но в том случае, когда рассматриваемые поля удовлетворяют некоторым динамическим уравнениям (в частности, волновым уравнениям) в *однородной* и *стационарной* среде, *свободной от источников* (т. е. сами уравнения однородны), то  $\omega$  и  $\mathbf{k}$  уже не независимы, а подчинены так называемому *дисперсионному уравнению*:

$$\Delta(\omega, \mathbf{k}) = 0, \quad (6.16)$$

Это уравнение следует из требования, чтобы плоская монохроматическая волна  $\exp(i\mathbf{k}\mathbf{r} - i\omega t)$  была *собственной* волной, т. е. была частным решением динамических уравнений.

Говоря на геометрическом языке, дисперсионное уравнение описывает в пространстве  $(\omega, \mathbf{k})$  некую трехмерную гиперповерхность — дисперсионную поверхность, — которая может быть и многолистной (например, в анизотропной среде). Вие точек этой гиперповерхности  $\omega\mathbf{k}$ -амплитуды полей равны нулю, а значит, равны нулю и спектральные плотности в (6.3) и (6.14). Другими словами,  $\xi^{(i)}(\omega, \mathbf{k})$  и  $G^{(i)}(\omega, \mathbf{k})$  содержат множителем дельта-функцию  $\delta[\Delta(\omega, \mathbf{k})]$  и, соответственно, снижается кратность интегралов Фурье: они распространяются фактически только на дисперсионную поверхность (всю или ее участки), т. е. на  $\omega$  и  $\mathbf{k}$ , удовлетворяющие дисперсионному уравнению. Уже в следующей главе мы непосредственно столкнемся с этим обстоятельством.

В тех случаях, когда стационарное поле  $\xi(t, \mathbf{r})$  является, как функция от  $t$ , *эргодическим* (ч. I, § 20), угловые скобки в предыдущих выражениях можно трактовать (в смысле вероятностной сходимости) как усреднение по времени, т. е. считать, например, что для произвольной детерминированной функции *практически* справедливо равенство

$$\langle f[\xi(t, \mathbf{r})] \rangle = \lim_{T \rightarrow \infty} \overline{f[\xi(t, \mathbf{r})]} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f[\xi(t, \mathbf{r})] dt, \quad (6.17)$$

причем для получения средних по ансамблю путем временного усреднения можно с достаточной точностью ограничиваться конечными  $T$ , существенно превышающими время корреляции  $\tau_{\text{кор}}$ :

$$T \gg \tau_{\text{кор}}. \quad (6.18)$$

Наряду с эргодичностью по времени, можно ввести понятие *пространственной* и *пространственно-временной* эргодичности. Так, для однородных пространственно эргодических полей средние по ансамблю «совпадают» в смысле сходимости по вероятности со средними по пространству. Практически это означает, что для произвольной детерминированной функции  $f$  можно считать справедливым равенство

$$\langle f[\xi(t, \mathbf{r})] \rangle = \lim_{V \rightarrow \infty} \overline{f_V[\xi(t, \mathbf{r})]} = \lim_{V \rightarrow \infty} \frac{1}{V} \int_V f[\xi(t, \mathbf{r})] d^3r, \quad (6.19)$$

где  $V$  — пространственная область, по которой ведется усреднение. Предельный переход  $V \rightarrow \infty$  опять-таки может быть приостановлен на областях  $V$ , поперечник [которых  $L$  велик по сравнению с радиусом корреляции  $l_\xi$ , т. е. выполняется

неравенство, аналогичное (6.18):

$$L \sim V^{1/2} \gg l_{\xi}. \quad (6.20)$$

О пространственно-временной эргодичности стационарных и однородных полей говорят тогда, когда имеет место сходимость по вероятности как при  $T \rightarrow \infty$ , так и при  $V \rightarrow \infty$ , что практически означает одновременное выполнение равенств (6.17) и (6.19). Возможны также случаи, когда поле  $\xi(t, \mathbf{r})$  является эргодическим только по части пространственных аргументов, например только в плоскости  $(x, y)$  или только на поверхности сферы.

Наконец, иногда удобно пользоваться понятием квазиэргодических полей, которые находятся по отношению к эргодическим полям в таком же положении, как квазигомогенные поля по отношению к однородным. Иными словами, эти поля являются эргодическими лишь в объемах, малых по сравнению с характеристическими масштабами  $L_{\xi}$  изменения статистических характеристик поля (среднего значения, дисперсии и т. д.). В отличие от (6.19), область пространственного усреднения для квазиэргодических полей должна быть ограничена сверху масштабом  $L_{\xi}$ , но при этом должно по-прежнему выполняться неравенство (6.20). Следовательно, о квазиэргодичности полей можно говорить лишь при таких условиях, когда можно ввести объем усреднения  $V$ , который удовлетворял бы двухстороннему неравенству

$$L_{\xi} \gg V^{1/2} \gg l_{\xi}. \quad (6.21)$$

В феноменологической физике  $V$  называют обычно «физически бесконечно малым объемом». Этот объем должен быть, с одной стороны, достаточно малым, чтобы в его пределах исследуемые поля были статистически однородными (в механике и электродинамике сплошных сред обычно требуется лишь постоянство в объеме  $V$  средних полей), а с другой — настолько большим, чтобы в пределах  $V$  поле  $\xi$  испытывало достаточно много пространственных флуктуаций.

В физике сплошных сред под  $l_{\xi}$  следует понимать среднее расстояние между источниками возмущений, которыми могут быть отдельные молекулы, дислокации, вкрапления и т. д. Если  $n$  — концентрация молекул (или других возмущающих объектов), то  $l_{\xi} \sim n^{-1/3}$ , и тогда неравенство  $V^{1/2} \gg l_{\xi}$  принимает вид  $nV \gg 1$ , что и отвечает большому числу молекул в физически бесконечно малом объеме  $V$ . При выполнении этого неравенства усреднение по ансамблю молекул можно заменять усреднением по малой пространственной области  $V$ . Таким образом, в макроскопической физике существенно используется предположение о пространственной квазиэргодичности тех или иных полей, характеризующих состояние сплошных сред.



## § 7. Функциональный метод описания случайных полей

Как мы знаем, полное описание вещественной случайной величины  $\xi$  дается ее характеристической функцией  $\Phi_{\xi}(u) = \langle \exp(iu\xi) \rangle$ , при помощи которой можно найти как плотность вероятностей  $\xi$ , так и моменты разного порядка, а также кумулянты.

Для  $N$ -мерной случайной величины  $\xi = \{\xi_1, \dots, \xi_N\}$  характеристическая функция имеет вид  $\Phi_{\xi}(u) = \langle \exp(iu\xi) \rangle$ , где  $u\xi = \sum_{j=1}^N u_j \xi_j$  и тоже содержит всю информацию о совокупности случайных величин  $\xi_j$  (с. I, § 9).

При рассмотрении общих вопросов, обсуждаемых в данном параграфе, целесообразно отказаться на время от ограничения числа параметров, от которых зависят интересующие нас функции, и вернуться от четырех координат  $x, y, z, t$  к  $n$ -мерному пространству параметров  $x = \{x_1, \dots, x_n\}$ . Пусть  $\xi(x_1, \dots, x_n) = \xi(x)$  — вещественное случайное поле в этом пространстве (в частном случае единственного параметра,  $n=1$ , мы будем, как и ранее, обычно считать, что параметр  $x$  является временем  $t$ ).

Рассмотрим теперь другой, по существу равносильный, но во многих задачах более удобный и лаконичный способ задания случайных функций — при помощи так называемых *характеристических функционалов*. Убедимся прежде всего, что полное вероятностное описание поля  $\xi(x)$  дается *характеристическим функционалом*

$$\Phi[u] = \langle \exp \left\{ i \int \xi(x) u(x) dx \right\} \rangle, \quad dx = dx_1 \dots dx_n. \quad (7.1)$$

Здесь  $u(x)$  — произвольная функция, для которой интеграл в (7.1) сходится при всех возможных реализациях  $\xi(x)$ . Функция  $u(x)$  заменяет дискретную совокупность переменных  $u = \{u_1, \dots, u_N\}$ , от которых зависит характеристическая функция для  $N$ -мерной случайной величины  $\xi = \{\xi_1, \dots, \xi_N\}$ . Подчеркнем, что функционал  $\Phi[u]$  должен быть известен для *любой* функции  $u(x)$ .

Зная  $\Phi[u]$ , можно получить характеристическую функцию для любой совокупности случайных величин  $\xi_i = \xi(x_i)$ ,  $\dots$ ,  $\xi_N = \xi(x_N)$ . Для этого достаточно взять в качестве аргумента  $u(x)$  функционала  $\Phi[u]$  функцию вида

$$u_N(x) = u_1 \delta(x - x_1) + \dots + u_N \delta(x - x_N).$$

Тогда

$$\int \xi(x) u_N(x) dx = \sum_1^N u_j \xi(x_j)$$

и функционал  $\Phi[u]$  превращается в характеристическую функцию  $N$ -мерной случайной величины  $\{\xi_1, \dots, \xi_N\}$ , т. е. значений поля  $\xi(x)$  в  $N$  точках  $x_1, \dots, x_N$ :

$$\Phi[u_N] = \left\langle \exp \left\{ i \sum_1^N u_j \xi(x_j) \right\} \right\rangle = \varphi(u_1, \dots, u_N).$$

Но характеристическая функция  $\varphi(u_1, \dots, u_N)$  позволяет, как мы знаем, найти  $N$ -мерную плотность вероятностей  $w_N(\xi_1, \dots, \xi_N)$ . Тем самым, ясно, что характеристический функционал  $\Phi[u]$  *однозначно* определяет плотности  $w_N(\xi_1, \dots, \xi_N)$  при *любом*  $N$  и при *любом* выборе точек  $x_1, \dots, x_N$ , т. е. полностью определяет случайное поле  $\xi(x)$ .

Зная функционал  $\Phi[u]$ , мы можем найти также моменты (в частности, корреляционные функции) и другие статистические характеристики случайного поля  $\xi(x)$ . Однако для этого необходимо уметь дифференцировать функционал  $\Phi[u]$  по *функциональному* аргументу  $u(x)$ . Поэтому часть этого параграфа будет посвящена чисто математическим вопросам, касающимся функционалов.

Прежде всего дадим *общее определение* функционала. Мы говорим, что функционал задан, если установлено правило, по которому *каждой функции, принадлежащей области определения функционала, сопоставлено число, называемое значением функционала на этой функции*. Приведем некоторые примеры функционалов:

$$F_1[u] = \int A(x) u(x) dx,$$

$$F_N[u] = \int A_N(x_1, \dots, x_N) u(x_1) \dots u(x_N) dx_1 \dots dx_N,$$

$$\Phi_G[u] = \exp \left\{ -1/2 \iint \psi(x_1, x_2) u(x_1) u(x_2) dx_1 dx_2 \right\}.$$

Здесь  $F_1[u]$  — линейный функционал,  $F_N[u]$  функционал  $N$ -й степени,  $\Phi_G[u]$  — гауссов функционал (ниже будет показано, что  $\Phi_G[u]$  — характеристический функционал для гауссовых случайных полей). Отметим, что функционал зависит от *вида функции*  $u(x)$  в целом (т. е. от ее поведения во *всем* пространстве  $x$ ), в силу чего в обозначении  $F[u]$  мы не пишем явно зависимости  $u$  от ее аргумента  $x$ . В тех случаях, когда на эту зависимость необходимо указывать, мы будем отмечать аргумент функции  $u$  тильдой:  $F[u(\tilde{x})]$ . Однако при такой записи следует иметь в виду, что сам функционал  $F$  не зависит от  $\tilde{x}$  (подобно тому как в тензорной алгебре величина  $x_k u_k$  не зависит от «немого» индекса  $k$  — индекса суммирования).

Рассмотрим значения одного и того же функционала  $F$  для двух различных «значений» его аргумента, а именно для функций

$u(x)$  и  $u(x) + \delta u(x)$ . Будем при этом считать, что функция  $\delta u(x)$  отлична от нуля лишь в некоторой области  $\Delta x$  вокруг точки  $x_0$ . *Функциональной* (или *вариационной*) *производной* функционала  $F[u]$  в точке  $x_0$  называется предел

$$\frac{\delta F[u]}{\delta u[x_0]} = \lim_{\substack{|\Delta x| \rightarrow 0 \\ \max |\delta u| \rightarrow 0}} \frac{F[u + \delta u] - F[u]}{\int_{\Delta x} \delta u(x) dx}, \quad (7.2)$$

конечно, при условиях, что этот предел существует и не зависит ни от вида  $\delta u(x)$ , ни от способа стягивания к нулю объема области  $|\Delta x|$ , ни от закона, по которому стремится к нулю максимум модуля функции  $\delta u(x)$ . Приведем некоторые примеры.

Найдем функциональную производную линейного функционала  $F_1[u] = \int A(x') u(x') dx'$  в точке  $x_0$ . Имеем

$$\begin{aligned} F_1[u + \delta u] &= \int A(x') [u(x') + \delta u(x')] dx', \\ F_1[u + \delta u] - F_1[u] &= \int_{\Delta x} A(x') \delta u(x') dx', \end{aligned}$$

так что

$$\frac{F_1[u + \delta u] - F_1[u]}{\int_{\Delta x} \delta u(x') dx'} = \frac{\int_{\Delta x} A(x') \delta u(x') dx'}{\int_{\Delta x} \delta u(x') dx'}.$$

Применяя теорему о среднем к интегралу в числителе, находим

$$\int_{\Delta x} A(x') \delta u(x') dx' = A(x_1) \int_{\Delta x} \delta u(x') dx',$$

где точка  $x_1$  принадлежит  $\Delta$ -окрестности точки  $x_0$ . Переходя к пределу при  $|\Delta x| \rightarrow 0$ , получаем

$$\frac{\delta}{\delta u(x_0)} \int A(x') u(x') dx' = A(x_0). \quad (7.3)$$

Таким же образом производится дифференцирование и более сложных функционалов (см. задачу 13).

Выберем в качестве  $\delta u(x)$  функцию  $\delta u(x) = \lambda \delta(x - x_0)$  (как и ранее,  $\delta(x - x_0) = \delta(x_1 - x_{01}) \dots \delta(x_n - x_{0n})$ ). Тогда  $\int \delta u(x') dx' = \lambda$ , согласно (7.2),

$$\begin{aligned} \frac{\delta F[u]}{\delta u(x_0)} &= \lim_{\lambda \rightarrow 0} \frac{F[u(\bar{x}) + \lambda \delta(\bar{x} - x_0)] - F[u(\bar{x})]}{\lambda} = \\ &= \left\{ \frac{\partial}{\partial \lambda} F[u(\bar{x}) + \lambda \delta(\bar{x} - x_0)] \right\}_{\lambda=0}. \quad (7.4) \end{aligned}$$

Формула (7.4) сводит операцию вычисления функциональной производной к обычному дифференцированию.

Приведем теперь формальные правила функционального дифференцирования, аналогичные обычным правилам.

Если  $A_1$  и  $A_2$  не зависят от  $u$ , то

$$\frac{\delta}{\delta u(x)} \{A_1 \Phi_1[u] + A_2 \Phi_2[u]\} = A_1 \frac{\delta \Phi_1[u]}{\delta u(x)} + A_2 \frac{\delta \Phi_2[u]}{\delta u(x)}. \quad (7.5)$$

Далее,

$$\frac{\delta}{\delta u(x)} \{\Phi_1[u] \Phi_2[u]\} = \frac{\delta \Phi_1[u]}{\delta u(x)} \Phi_2[u] + \Phi_1[u] \frac{\delta \Phi_2[u]}{\delta u(x)}. \quad (7.6)$$

Пусть  $f(z)$  — функция от  $z$ . Рассмотрим функционал вида  $\Phi[u] = f(F[u])$  (функция от функционала). Тогда

$$\frac{\delta}{\delta u(x)} f(F[u]) = f'(F[u]) \frac{\delta F[u]}{\delta u(x)}, \quad (7.7)$$

где  $f' = \frac{df(z)}{dz} \Big|_{z=F[u]}$  — обычная производная функции  $f$  по ее аргументу. Формулы (7.5)–(7.7) непосредственно следуют из (7.4). Наконец, положив в (7.3)  $A(x') = \delta(x' - x_1)$ , получаем

$$\frac{\delta u(x_1)}{\delta u(x_0)} = \delta(x_1 - x_0). \quad (7.8)$$

Формулы (7.5)–(7.8) позволяют, не прибегая к формальному определению (7.2) операции функционального дифференцирования, дифференцировать большинство функционалов, с которыми приходится встречаться в конкретных задачах. В частности, из (7.5) следует, что можно производить дифференцирование функционалов под знаком интеграла или производной.

В качестве еще одного примера найдем функциональную производную от гауссова функционала:

$$\frac{\delta \Phi_G[u]}{\delta u(x)} = \frac{\delta}{\delta u(x)} \exp \left\{ -\frac{1}{2} \iint \Psi(x_1, x_2) u(x_1) u(x_2) dx_1 dx_2 \right\}.$$

В соответствии с (7.7) получаем

$$\frac{\delta \Phi_G[u]}{\delta u(x)} = \Phi_G[u] \left\{ -\frac{1}{2} \frac{\delta}{\delta u(x)} \iint \Psi(x_1, x_2) u(x_1) u(x_2) dx_1 dx_2 \right\}.$$

На основании (7.5) дифференцируем под знаком интеграла, используя при этом формулы (7.6) и (7.8):

$$\begin{aligned} \frac{\delta \Phi_G[u]}{\delta u(x)} &= -\frac{1}{2} \Phi_G[u] \iint \Psi(x_1, x_2) [\delta(x_1 - x) u(x_2) + \\ &\quad + u(x_1) \delta(x_2 - x)] dx_1 dx_2 = \\ &= -\frac{1}{2} \Phi_G[u] \int [\Psi(x, x') + \Psi(x', x)] u(x') dx'. \end{aligned}$$

В данном примере можно считать, что  $\psi(x_1, x_2) = \psi(x_2, x_1)$ , так как несимметричная часть  $\psi(x_1, x_2)$  не вносит вклада в функционал, поскольку интеграл от произведения несимметричной функции на симметричную  $u(x_1)$  и  $u(x_2)$  равен нулю. Поэтому полученную производную можно записать короче:

$$\frac{\delta \Phi_G[u]}{\delta u(x)} = -\Phi_G[u] \int \psi(x, x') u(x') dx'.$$

Функционал  $F[u+v]$  можно разложить в функциональный ряд Тейлора по  $v(\tilde{x})$  (см., например, [1], приложение I):

$$F[u+v] = F[u] + \int dx v(x) \frac{\delta F[u]}{\delta u(x)} + \\ + \frac{1}{2!} \iint dx_1 dx_2 v(x_1) v(x_2) \frac{\delta^2 F[u]}{\delta u(x_1) \delta u(x_2)} + \dots \quad (7.9)$$

Подобно операторной записи формулы  $f(x+\xi) = \exp\left\{\xi \frac{\partial}{\partial x}\right\} f(x)$  для обычного ряда Тейлора, удобно записать и формулу (7.9) в операторном виде:

$$F[u+v] = \left\{ 1 + \int dx v(x) \frac{\delta}{\delta u(x)} + \right. \\ \left. + \frac{1}{2!} \iint dx_1 dx_2 v(x_1) v(x_2) \frac{\delta}{\delta u(x_1)} \frac{\delta}{\delta u(x_2)} + \dots \right\} F[u] = \\ = \left\{ 1 + \int dx v(x) \frac{\delta}{\delta u(x)} + \frac{1}{2!} \left[ \int dx v(x) \frac{\delta}{\delta u(x)} \right]^2 + \dots \right\} F[u].$$

В фигурных скобках мы имеем степенное разложение экспоненты, так что

$$F[u+v] = \exp \left\{ \int dx v(x) \frac{\delta}{\delta u(x)} \right\} F[u]. \quad (7.10)$$

Вернемся к характеристическому функционалу (7.1) и подействуем на него оператором  $\frac{1}{i} \frac{\delta}{\delta u(x)}$ . Выполняя дифференцирование под знаком среднего, находим

$$\frac{1}{i} \frac{\delta \Phi[u]}{\delta u(x)} = \langle \xi(x) \exp \left\{ i \int u(x') \xi(x') dx' \right\} \rangle. \quad (7.11)$$

Таким образом, оператор  $\frac{1}{i} \frac{\delta}{\delta u(x)}$ , действуя на  $\Phi[u]$ , вводит под знак усреднения множитель  $\xi(x')$ . Повторяя эту операцию, приходим к формуле

$$\frac{1}{i} \frac{\delta}{\delta u(x_1)} \dots \frac{1}{i} \frac{\delta}{\delta u(x_N)} \Phi[u] = \\ = \langle \xi(x_1) \dots \xi(x_N) \exp \left\{ i \int u(x') \xi(x') dx' \right\} \rangle. \quad (7.12)$$

Если теперь положить здесь  $u=0$ , то получим

$$\langle \xi(x_1) \dots \xi(x_N) \rangle = \frac{1}{i^N} \frac{\delta^N \Phi[u]}{\delta u(x_1) \dots \delta u(x_N)} \Big|_{u(\tilde{x})=0}. \quad (7.13)$$

Таким образом, моменты (7.13) случайного поля  $\xi(x)$  можно находить как значения функциональных производных характеристического функционала при нулевом значении его аргумента.

Если разложить экспоненту в (7.1) в ряд и записать  $k$ -ю степень интеграла в виде  $k$ -кратного интеграла, то получится формула

$$\Phi[u] = 1 + \sum_1^{\infty} \frac{i^k}{k!} \int \langle \xi(x_1) \dots \xi(x_k) \rangle u(x_1) \dots u(x_k) dx_1 \dots dx_k, \quad (7.14)$$

представляющая собой разложение характеристического функционала  $\Phi[u]$  в функциональный ряд Тейлора.

Рассмотрим теперь функционал  $\Theta[u] = \ln \Phi[u]$ , так что  $\Phi[u] = \exp \{\Theta[u]\}$ . Разложение  $\Theta[u]$  в ряд Тейлора определяет *кумулянтные* (или корреляционные) функции различных порядков:

$$\Theta[u] = \sum_1^{\infty} \frac{i^k}{k!} \int \psi_k(x_1, \dots, x_k) u(x_1) \dots u(x_k) dx_1 \dots dx_k. \quad (7.15)$$

Кумулянтная функция  $k$ -го порядка ( $k=1, 2, \dots$ ) определена формулой

$$\psi_k(x_1, \dots, x_k) = \frac{1}{i^k} \frac{\delta^k \ln \Phi[u]}{\delta u(x_1) \dots \delta u(x_k)} \Big|_{u=0}, \quad (7.16)$$

аналогичной формуле для кумулянтов случайной величины. Из (7.13) и (7.16) вытекает связь между моментными и кумулянтными функциями (см. задачу 16). В частности,

$$\begin{aligned} \psi_1(x) &= \overline{\xi(x)}, \\ \psi_2(x_1, x_2) &= \psi(x_1, x_2) = \langle \xi(x_1) \xi(x_2) \rangle - \overline{\xi(x_1)} \overline{\xi(x_2)}. \end{aligned} \quad (7.17)$$

Выше мы рассмотрели несколько примеров функционалов общего вида. Обратимся теперь к примерам *характеристических функционалов*.

1) Пусть  $\xi(x)$  — нормальное случайное поле. Обозначим через  $\mathcal{L}$  интеграл

$$\mathcal{L} = \int \xi(x) u(x) dx.$$

Тогда характеристический функционал  $\Phi_G[u]$  запишется в виде

$$\Phi_G[u] = \langle \exp(i\mathcal{L}) \rangle.$$

Интеграл от гауссовой случайной функции является гауссовой случайной величиной, что позволяет сразу же написать плотность вероятностей для  $\mathcal{L}$ :

$$w(\mathcal{L}) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(\mathcal{L} - \bar{\mathcal{L}})^2}{2\sigma^2}\right).$$

Из выражения для  $\mathcal{L}$  очевидно, что параметры  $\bar{\mathcal{L}}$  и  $\sigma^2 = \langle (\mathcal{L} - \bar{\mathcal{L}})^2 \rangle = \langle (\tilde{\mathcal{L}})^2 \rangle$  этого распределения выражаются следующим образом:

$$\begin{aligned}\bar{\mathcal{L}} &= \int \bar{\xi}(\mathbf{x}) u(\mathbf{x}) d\mathbf{x}, \\ \sigma^2 &= \iint \psi_{\xi}(x_1, x_2) u(x_1) u(x_2) dx_1 dx_2.\end{aligned}$$

Вычисляя теперь  $\langle \exp(i\mathcal{L}) \rangle$ , получаем известный результат:

$$\begin{aligned}\Phi_G[u] &= \int e^{i\mathcal{L}} w(\mathcal{L}) d\mathcal{L} = \\ &= \frac{1}{\sqrt{2\pi\sigma}} \int \exp\left(i\mathcal{L} - \frac{1}{2\sigma^2} (\mathcal{L} - \bar{\mathcal{L}})^2\right) d\mathcal{L} = \exp\left(i\bar{\mathcal{L}} - \frac{1}{2}\sigma^2\right),\end{aligned}$$

а подстановка выражений для  $\bar{\mathcal{L}}$  и  $\sigma^2$  дает

$$\begin{aligned}\Phi_G[u] &= \exp\left\{i \int \bar{\xi}(\mathbf{x}) u(\mathbf{x}) d\mathbf{x} - \right. \\ &\quad \left. - \frac{1}{2} \iint \psi_{\xi}(x_1, x_2) u(x_1) u(x_2) dx_1 dx_2\right\}. \quad (7.18)\end{aligned}$$

Отсюда следует, что

$$\begin{aligned}\Theta_G[u] = \ln \Phi_G[u] &= \\ &= i \int \bar{\xi}(\mathbf{x}) u(\mathbf{x}) d\mathbf{x} - \frac{1}{2} \iint \psi_{\xi}(x_1, x_2) u(x_1) u(x_2) dx_1 dx_2.\end{aligned} \quad (7.19)$$

Сопоставление (7.19) с (7.15) показывает, что для гауссовых случайных полей все кумулянтные функции выше второго порядка равны нулю.

Центральные моменты нормального случайного поля обладают следующими свойствами (ч. I, задача 6 к гл. II):

$$\begin{aligned}\langle \tilde{\xi}_1 \dots \tilde{\xi}_{2k+1} \rangle &= 0, \quad \tilde{\xi}_i \equiv \tilde{\xi}(x_i), \\ \langle \tilde{\xi}_1 \dots \tilde{\xi}_{2k} \rangle &= \sum_{p. n.} \langle \tilde{\xi}_{\alpha_1} \tilde{\xi}_{\alpha_2} \rangle \dots \langle \tilde{\xi}_{\alpha_{2k-1}} \tilde{\xi}_{\alpha_{2k}} \rangle,\end{aligned} \quad (7.20)$$

где р. н. под символом  $\sum$  означает, что суммирование производится по всем возможным разбиениям индексов 1, 2, ..., 2k на пары. Нетрудно подсчитать, что число слагаемых в сумме, входящей в (7.20), равно  $(2k-1)!! = 1 \cdot 3 \cdot 5 \dots (2k-1)$ . Например, при  $k=2$

$$\langle \tilde{\xi}_1 \tilde{\xi}_2 \tilde{\xi}_3 \tilde{\xi}_4 \rangle = \langle \tilde{\xi}_1 \tilde{\xi}_2 \rangle \langle \tilde{\xi}_3 \tilde{\xi}_4 \rangle + \langle \tilde{\xi}_1 \tilde{\xi}_3 \rangle \langle \tilde{\xi}_2 \tilde{\xi}_4 \rangle + \langle \tilde{\xi}_1 \tilde{\xi}_4 \rangle \langle \tilde{\xi}_2 \tilde{\xi}_3 \rangle.$$

Аналогичная формула для шестого момента содержит 15 слагаемых. Определения гауссова поля (7.18) и (7.20) равносильны (см. задачу 17).

2) Пуассоновское случайное поле строится следующим образом. Рассмотрим сначала конечную область  $V$  пространства параметров  $\mathbf{x} = \{x_1, \dots, x_n\}$ . Пусть  $m$  — случайная величина, распределенная по закону Пуассона  $P(m) = (\bar{m})^m e^{-\bar{m}} / m!$ , где  $\bar{m}$  — среднее значение  $m$ . Выберем в  $V$  случайным образом  $m$  точек  $x_1, \dots, x_m$ , считая, что плотность вероятностей для координат каждой из этих точек равна  $V^{-1}$  при  $x_k \in V$  и нулю вне  $V$  и что координаты всех  $m$  точек статистически независимы. Возьмем, далее,  $m$  случайных величин  $A_1, \dots, A_m$ , взаимно независимых и имеющих одинаковую плотность вероятностей  $\omega(A)$ , а значит, и характеристическую функцию  $\chi(\lambda)$ . Пусть  $g(\mathbf{x})$  — фиксированная детерминированная функция. Пуассоновское случайное поле мы определим теперь формулой

$$\xi(\mathbf{x}) = \sum_{k=0}^m A_k g(\mathbf{x} - \mathbf{x}_k). \quad (7.21)$$

Согласно построению в (7.21) содержатся следующие случайные параметры: «амплитуды»  $A_k$ , положения «центров»  $\mathbf{x}_k$  и число слагаемых  $m$ . Мы распространяем здесь на случай поля в  $n$ -мерном пространстве параметров  $\mathbf{x}$  то, что в ч. I было сделано для однопараметрических случайных функций, т. е. для случайных процессов (см., в частности, ч. I, гл. II).

Характеристический функционал случайного поля (7.21) имеет вид (см. задачу 18)

$$\Phi[u] = \exp \left\{ \nu \int \left[ \chi \left( \int g(\mathbf{x} - \mathbf{x}') u(\mathbf{x}) d\mathbf{x} \right) - 1 \right] d\mathbf{x}' \right\}, \quad (7.22)$$

где  $\nu = \bar{m}/V$  — средняя плотность случайных точек. Область  $V$  в (7.22) можно считать и неограниченной ( $V \rightarrow \infty$ ,  $\bar{m} \rightarrow \infty$  при  $\nu = \bar{m}/V = \text{const}$ ).

Из (7.22) нетрудно вывести кумулянтные функции пуассоновского случайного поля (см. задачу 19). Они имеют вид

$$\Phi_k(x_1, \dots, x_k) = \nu \langle A^k \rangle \int g(x_1 - \mathbf{x}) g(x_2 - \mathbf{x}) \dots g(x_k - \mathbf{x}) d\mathbf{x}. \quad (7.23)$$

Гауссово случайное поле можно получить из пуассоновского, если определенным образом устремить в бесконечность  $\nu$ , уменьшая в то же время «амплитуды»  $A$  (см. задачу 20).

При решении многих физических задач возникает необходимость вычисления величин вида  $\langle \xi(\mathbf{x}) R[\xi] \rangle$ , где  $R[\xi]$  — некоторый функционал от случайного поля  $\xi(\mathbf{x})$ . Выведем, следуя работе [5], полезную формулу, позволяющую находить эту величину.



Рассмотрим среднее значение выражения

$$\langle \xi(x) R[\xi + \eta] \rangle \equiv f[x, \eta],$$

где  $\eta = \eta(\tilde{x})$  — некоторая *детерминированная* функция. Воспользовавшись формулой (7.10) для ряда Тейлора, представим функционал  $R[\xi + \eta]$  в виде

$$R[\eta + \xi] = \exp \left\{ \int dx' \xi(x') \frac{\delta}{\delta \eta(x')} \right\} R[\eta].$$

Тогда

$$f[x, \eta] = \langle \xi(x) \exp \left\{ \int dx' \xi(x') \frac{\delta}{\delta \eta(x')} \right\} R[\eta] \rangle. \quad (7.24)$$

Разделив и умножив это равенство на оператор

$$\left\langle \exp \left\{ \int dx' \xi(x') \frac{\delta}{\delta \eta(x')} \right\} \right\rangle$$

(он коммутирует с входящим в (7.24) оператором), получаем

$$f[x, \eta] = \frac{\left\langle \exp \left\{ \int dx' \xi(x') \frac{\delta}{\delta \eta(x')} \right\} \xi(x) \right\rangle}{\left\langle \exp \left\{ \int dx' \xi(x') \frac{\delta}{\delta \eta(x')} \right\} \right\rangle} \times \\ \times \left\langle \exp \left\{ \int dx' \xi(x') \frac{\delta}{\delta \eta(x')} \right\} \right\rangle R[\eta]. \quad (7.25)$$

Введем функционал  $\Omega$ , зависящий от функции  $u(\tilde{x})$  и кроме того, являющийся обычной функцией параметра  $x$ :

$$\Omega[x, u] = \frac{\left\langle \xi(x) \exp \left\{ i \int dx' \xi(x') u(x') \right\} \right\rangle}{\left\langle \exp \left\{ i \int dx' \xi(x') u(x') \right\} \right\rangle}. \quad (7.26)$$

С его помощью можно записать (7.25) в виде

$$f[x, \eta] = \Omega \left[ x, \frac{1}{i} \frac{\delta}{\delta \eta} \right] \left\langle \exp \left\{ \int dx' \xi(x') \frac{\delta}{\delta \eta(x')} \right\} R[\eta] \right\rangle,$$

где в последнем сомножителе неслучайная величина  $R[\eta]$  введена под знак усреднения. Снова воспользовавшись формулой (7.10), получаем

$$f[x, \eta] = \Omega \left[ x, \frac{1}{i} \frac{\delta}{\delta \eta} \right] \langle R[\eta + \xi] \rangle,$$

или, внося неслучайный оператор под знак среднего,

$$f[x, \eta] = \left\langle \Omega \left[ x, \frac{1}{i} \frac{\delta}{\delta \eta} \right] R[\eta + \xi] \right\rangle.$$

Поскольку функционал  $R$  зависит лишь от суммы  $\eta + \xi$ , можно заменить оператор дифференцирования по  $\eta$  оператором дифференцирования по  $\xi$ . Таким образом,

$$f[x, \eta] = \langle \xi(x) R[\eta + \xi] \rangle = \left\langle \Omega \left[ x, \frac{1}{i} \frac{\delta}{\delta \xi} \right] R[\eta + \xi] \right\rangle,$$

причем здесь уже можно приравнять вспомогательную функцию  $\eta$  нулю. Это приводит к окончательной формуле, полученной В. И. Клячкиным [5]:

$$\langle \xi(x) R[\xi] \rangle = \left\langle \Omega \left[ x, \frac{1}{i} \frac{\delta}{\delta \xi} \right] R[\xi] \right\rangle. \quad (7.27)$$

Функционал  $\Omega[x, u]$  можно, согласно (7.26), (7.11) и (7.1), выразить через логарифм характеристического функционала поля  $\xi(x)$  следующим образом:

$$\Omega[x, u] = \frac{1}{\Phi[u]} \frac{1}{i} \frac{\delta \Phi[u]}{\delta u(x)} = \frac{1}{i} \frac{\delta \ln \Phi[u]}{\delta u(x)} = \frac{1}{i} \frac{\delta \Theta[u]}{\delta u(x)}. \quad (7.28)$$

Воспользуемся теперь разложением (7.15) функционала  $\Theta$  по кумулянтным функциям. Дифференцируя (7.15), получаем с учетом результата задачи 13

$$\begin{aligned} \Omega[x, u] &= \frac{1}{i} \frac{\delta \Theta[u]}{\delta u(x)} = \\ &= \sum_{k=1}^{\infty} \frac{i^{k-1}}{(k-1)!} \int \psi_k(x, x_1, \dots, x_{k-1}) u(x_1) \dots u(x_{k-1}) dx_1 \dots dx_{k-1}. \end{aligned}$$

Заменяя здесь  $u(x_i)$  на  $\frac{1}{i} \frac{\delta}{\delta \xi(x_i)}$  и подставив результат в (7.27), находим

$$\begin{aligned} \langle \xi(x) R[\xi] \rangle &= \sum_{k=1}^{\infty} \frac{1}{(k-1)!} \int \psi_k(x, x_1, \dots, x_{k-1}) \times \\ &\times \left\langle \frac{\delta^{(k-1)} R[\xi]}{\delta \xi(x_1) \dots \delta \xi(x_{k-1})} \right\rangle dx_1 \dots dx_{k-1}. \end{aligned} \quad (7.29)$$

Рассмотрим частный случай гауссова поля  $\xi(x)$ , у которого  $\psi_3 = \psi_4 = \dots = 0$  и, следовательно, в (7.29) остаются только два первых слагаемых с  $\psi_1(x) = \xi(x)$  и  $\psi_2(x, x') = \psi_{\xi}(x, x')$ :

$$\langle \xi(x) R[\xi] \rangle = \overline{\xi(x)} \langle R[\xi] \rangle + \int \psi_{\xi}(x, x') \left\langle \frac{\delta R[\xi]}{\delta \xi(x')} \right\rangle dx'. \quad (7.30)$$

Формула (7.30) была получена К. Фуруцу [6], Е. А. Новиковым [7] и М. Д. Донскером [8].

Укажем, как изменятся некоторые из предыдущих формул, если рассматривать совокупность  $N$  случайных полей ( $N$ -мерное поле)  $\xi_1(x), \dots, \xi_N(x)$ . Характеристический функционал определяется тогда формулой

$$\Phi[u_1, \dots, u_N] = \left\langle \exp \left\{ i \sum_{k=1}^N \int u_k(x) \xi_k(x) dx \right\} \right\rangle. \quad (7.31)$$

Формулы (7.5)–(7.7) сохраняют прежний вид, но с заменой в них  $\delta/\delta u(x)$  на  $\delta/\delta u_j(x)$ . Вместо (7.8) теперь справедлива формула

$$\frac{\delta u_i(x)}{\delta u_k(x')} = \delta_{ik} \delta(x - x'). \quad (7.32)$$

Формула Фуруцу—Новикова (7.30) принимает вид

$$\begin{aligned} \langle \xi_i(x) R[\xi_1, \dots, \xi_N] \rangle &= \\ &= \overline{\xi_i(x)} \langle R \rangle + \sum_{k=1}^N \int \langle \tilde{\xi}_i(x) \tilde{\xi}_k(x') \rangle \left\langle \frac{\delta R}{\delta \xi_k(x')} \right\rangle dx'. \end{aligned} \quad (7.33)$$

В заключение сформулируем свойства однородности и изотропности случайных полей на языке характеристического функционала. Статистическая однородность поля  $\xi(x)$  (в узком смысле) означает, что  $\langle \xi(x_1 + x_0) \dots \xi(x_k + x_0) \rangle = \langle \xi(x_1) \dots \xi(x_k) \rangle$ . Используя это равенство, можно записать разложение (7.14) в виде

$$\begin{aligned} \Phi[u(\tilde{x})] &= 1 + \sum_{k=1}^{\infty} \frac{i^k}{k!} \int \langle \xi(x_1 + x_0) \dots \xi(x_k + x_0) \rangle \times \\ &\quad \times u(x_1) \dots u(x_k) dx_1 \dots dx_k \end{aligned}$$

или, после замены переменных  $x_i + x_0 = x'_i$ , в виде

$$\begin{aligned} \Phi[u(\tilde{x})] &= 1 + \sum_{k=1}^{\infty} \frac{i^k}{k!} \int \langle \xi(x'_1) \dots \xi(x'_k) \rangle \times \\ &\quad \times u(x'_1 - x_0) \dots u(x'_k - x_0) dx'_1 \dots dx'_k. \end{aligned}$$

Но, как легко видеть, в правой части получился функционал  $\Phi[u(\tilde{x} - x_0)]$ . Таким образом, для статистически однородных полей

$$\Phi[u(\tilde{x})] = \Phi[u(\tilde{x} - x_0)], \quad (7.34)$$

т. е. характеристический функционал не меняется, если его аргумент  $u$  рассматривается в смещенной точке.

Аналогичным образом, если  $\hat{R}x = x'$  означает преобразование поворота в пространстве параметров, то статистическая

изотропность поля  $\xi(\mathbf{x})$  выражается равенством

$$\Phi[u(\tilde{\mathbf{x}})] = \Phi[u(\hat{R}^{-1}\tilde{\mathbf{x}})]. \quad (7.35)$$

Возможны случаи, когда поле статистически однородно или изотропно не по всем координатам  $x_1, \dots, x_n$ , а лишь в отношении некоторой их части, т. е. в некотором подпространстве параметров  $\mathbf{x}$ . Тогда инвариантность  $\Phi[u]$  должна иметь место при преобразованиях трансляции (вектор  $\mathbf{x}_0$  в (7.34)) или поворота (оператор  $\hat{R}$  в (7.35)), осуществляемых только в этом подпространстве.

Приведем в качестве примера характеристический функционал статистически однородного гауссова поля, у которого  $\bar{\xi} = 0$ . Для него

$$\Phi[u] = \exp \left\{ -\frac{1}{2} \iint \Psi_{\xi}(\mathbf{x}_1 - \mathbf{x}_2) u(\mathbf{x}_1) u(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2 \right\}. \quad (7.36)$$

Ясно, что (7.34) здесь выполнено. Если рассматриваемое поле статистически изотропно, то вместо (7.36) будет

$$\Phi[u] = \exp \left\{ -\frac{1}{2} \iint \Psi_{\xi}(|\mathbf{x}_1 - \mathbf{x}_2|) u(\mathbf{x}_1) u(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2 \right\}. \quad (7.37)$$

Как уже было сказано, этот параграф посвящен в основном правилам оперирования с функционалами и выводу ряда важных соотношений и формул. Примеры использования функционального метода при решении конкретных задач рассматриваются в некоторых из последующих глав. В задаче 22 функциональный метод применен для вывода приведенного в ч. I уравнения Эйнштейна—Фоккера (36.15) для вероятностей перехода случайного отклика дискретной динамической системы на гауссовы дельта-коррелированные по времени воздействия.

### Задачи

1. Найти пространственную спектральную плотность для анизотропного однородного случайного поля с гауссовой корреляционной функцией общего вида

$$\Psi_{\xi}(\mathbf{r}) = \sigma_{\xi}^2 \exp \left\{ -\frac{1}{2} \sum_{i,j} \alpha_{ij} x_i x_j \right\},$$

где  $x_1 = x$ ,  $x_2 = y$ ,  $x_3 = z$  и  $\text{Det} \|\alpha_{ij}\| \neq 0$ .

От а е т.

$$\Phi_{\xi}(\mathbf{k}) = \frac{\sigma_{\xi}^2}{\sqrt{(2\pi)^3 \text{Det} \|\alpha_{ij}\|}} \exp \left\{ -\frac{1}{2} \sum_{i,j} A_{ij} k_i k_j \right\},$$

где  $\|A_{ij}\| = \|\alpha_{ij}\|^{-1}$  — матрица, обратная матрице  $\|\alpha_{ij}\|$ . В частном случае, когда корреляционная функция имеет вид (2.29), пространственный спектр дается выражением (3.16).

2. Доказать следующее неравенство для статистических моментов порядка  $(n+1)$  комплексного поля  $\zeta(r)$ :

$$|B_{n,n}(1, \dots, 2n)|^2 \leq B_{n,n}(1, \dots, n; n, \dots, 1) B_{n,n}(n+1, \dots, 2n; 2n, \dots, n+1). \quad (1)$$

Решение. Неравенство следует из того, что при любом комплексном  $\lambda$

$$\langle |\lambda \zeta(1) \dots \zeta(n) + \zeta(n+1) \dots \zeta(2n)|^2 \rangle \geq 0.$$

Если положить

$$\lambda = \alpha [B_{n,n}(1, \dots, 2n)]^*,$$

где  $\alpha$  вещественно, то неравенство (1) представляет собой условие неотрицательности квадратичного по  $\alpha$  трехчлена. Частный случай (1) при  $n=1$  — это неравенство

$$|B(1, 2)|^2 \leq B(1, 1) B(2, 2),$$

аналогичное (2.20).

3. Если провести в пространстве какую-либо прямую линию и рассматривать значения однородного и изотропного поля  $\zeta(r)$  только на этой прямой, то получится случайная функция одной пространственной переменной — координаты, отсчитываемой вдоль этой прямой. Обозначим ее через  $z$ . Можно записать разложение корреляционной функции  $\psi_\zeta(z)$  в одномерный интеграл Фурье:

$$\psi_\zeta(z) = \int_{-\infty}^{\infty} \gamma_\zeta(k) \cos kz \, dk,$$

где  $\gamma_\zeta(k)$  — одномерная спектральная плотность, выражающаяся через  $\psi_\zeta(z)$  посредством обратного преобразования Фурье:

$$\gamma_\zeta(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi_\zeta(z) \cos kz \, dz. \quad (1)$$

Показать, что трехмерный пространственный спектр  $\Phi_\zeta(k)$  однородного изотропного поля  $\zeta(r)$  связан с одномерным спектром  $\gamma_\zeta(k)$  соотношением

$$\Phi_\zeta(k) = -\frac{1}{2\pi k} \frac{\partial \gamma_\zeta(k)}{\partial k}. \quad (2)$$

Указание. Надо продифференцировать (1) по  $k$  и сравнить результат с формулой (3.10).

Формула (2) удобна для расчета трехмерного спектра в тех довольно многочисленных случаях, когда интеграл (1) для  $\gamma_\zeta(k)$  вычисляется проще, чем интеграл (3.10) для  $\Phi_\zeta(k)$ . Примерами могут служить даже простые корреляционные функции (2.24) и (2.25).

Формула (2) используется также для нахождения трехмерного спектра случайного изотропного поля по одномерному спектру  $\gamma_\zeta(k)$ , получаемому по экспериментально измеренной одномерной корреляционной функции  $\psi_\zeta(z)$ .

4. Пусть, как и в предыдущей задаче,  $z$  — координата, отсчитываемая вдоль какого-либо выбранного направления в пространстве, а случайное поле  $\zeta(r)$  однородно в плоскости  $z = \text{const}$ . Вводя в плоскости  $z = \text{const}$  двумерный вектор  $\rho = (x, y)$ , так что  $r = (\rho, z)$ , можно записать корреляционную

функцию поля  $\zeta(\mathbf{r})$  в виде

$$\psi_{\zeta}(\rho, z_1 - z_2) = \langle \tilde{\zeta}(\rho_1, z_1) \tilde{\zeta}^*(\rho_1 - \rho, z_2) \rangle.$$

Эту функцию можно разложить в двумерный интеграл Фурье:

$$\psi_{\zeta}(\rho, z_1 - z_2) = \int_{-\infty}^{\infty} F_{\zeta}(\mathbf{k}_{\perp}, z_1 - z_2) \exp(i\mathbf{k}_{\perp} \rho) d^2 k_{\perp}, \quad (1)$$

где  $\mathbf{k}_{\perp} = (k_x, k_y)$ . Функция  $F_{\zeta}(\mathbf{k}_{\perp}, z_1 - z_2)$  называется *двумерной спектральной плотностью* поля или просто *двумерным* (пространственным) *спектром*. Показать, что для однородного и *изотропного* случайного поля двумерный спектр  $F_{\zeta}(\mathbf{k}_{\perp}, z_1 - z_2)$  и трехмерный спектр  $\Phi_{\zeta}(k)$  связаны между собой соотношениями

$$F_{\zeta}(\mathbf{k}_{\perp}, z_1 - z_2) = \int_{-\infty}^{\infty} \Phi_{\zeta}(\sqrt{k_z^2 + k_{\perp}^2}) \cos k_z(z_1 - z_2) dk_z, \quad (2)$$

$$\Phi_{\zeta}(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F_{\zeta}(\mathbf{k}_{\perp}, z_1 - z_2) d(z_1 - z_2). \quad (3)$$

**Решение.** Формула (2) следует из сопоставления (1) с трехмерным разложением (3.5), если учесть вещественность корреляционной функции  $\psi_{\zeta}(r)$  для изотропного поля. Обращая (2), находим

$$\Phi_{\zeta}(\sqrt{k_z^2 + k_{\perp}^2}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F_{\zeta}(\mathbf{k}_{\perp}, z_1 - z_2) \cos k_z z d(z_1 - z_2).$$

Формула (3) получается отсюда при  $k_z = 0$ . Отметим, что  $\Phi_{\zeta}(\sqrt{k_z^2 + k_{\perp}^2})$  как функция от  $k_z$  имеет характерный масштаб  $k_{\perp}$ . В силу этого из (2) следует, что  $F_{\zeta}(\mathbf{k}_{\perp}, z_1 - z_2)$  сосредоточена в области

$$k_{\perp} |z_1 - z_2| < 2\pi. \quad (4)$$

5. Выразить спектральное разложение корреляционной функции поля, изотропного в плоскости  $z = \text{const}$ , в виде однократного интеграла.

**Решение.** В общем случае корреляционная функция  $\psi_{\zeta}(x, y) = \psi_{\zeta}(\rho)$  и ее двумерный спектр  $F_{\zeta}(x_x, x_y) = F_{\zeta}(x)$  связаны прямым и обратным преобразованиями Фурье:

$$\psi_{\zeta}(\rho) = \int F_{\zeta}(x) \exp(ix\rho) d^2 x, \quad F_{\zeta}(x) = \frac{1}{(2\pi)^2} \int \psi_{\zeta}(\rho) \exp(-ix\rho) d^2 \rho. \quad (1)$$

В случае изотропного поля имеем  $\psi_{\zeta}(\rho) = \psi_{\zeta}(\rho)$  и  $F_{\zeta}(x) = F_{\zeta}(x)$ . Переход в формулах (1) к полярным координатам (соответственно в  $\rho$ -и  $x$ -пространствах) с последующим интегрированием по угловым переменным приводит к прямому и обратному преобразованиям Ганкеля:

$$\psi_{\zeta}(\rho) = 2\pi \int_0^{\infty} F_{\zeta}(x) J_0(x\rho) x dx, \quad F_{\zeta}(x) = \frac{1}{2\pi} \int_0^{\infty} \psi_{\zeta}(\rho) J_0(x\rho) \rho d\rho, \quad (2)$$

где  $\rho = \sqrt{x^2 + y^2}$ ,  $x = \sqrt{x_x^2 + x_y^2}$ , а  $J_0$  — нулевая бесселева функция.

6. Найти частотный спектр  $g_{\zeta}(\omega)$  замороженного поля  $\zeta(t, r)$ , если соответствующее покоящееся случайное поле  $v(r)$  обладает гауссовой пространственной корреляционной функцией  $\psi_v(r) = \sigma_v^2 \exp\{-r^2/2l^2\}$ .

Ответ.

$$g_{\zeta}(\omega) = \frac{\sigma_v^2 l}{\sqrt{2\pi} v} \exp\left\{-\frac{1}{2} \left(\frac{\omega l}{v}\right)^2\right\}.$$

7. Найти корреляционную функцию, пространственно-временной, пространственный и частотный спектры замороженного поля  $\zeta(t, r)$  при релятивистских значениях скорости  $v$ , если известна функция корреляции  $\psi_v(r)$  (или пространственный спектр  $\Phi_v(k)$ ) соответствующего покоящегося поля  $v(r)$ .

Ответ. Если скорость  $v$  направлена по оси  $x$ , то в соответствии с преобразованием Лоренца

$$\Psi_{\zeta}(\tau, r) = \psi_v\left(\frac{x - v\tau}{\sqrt{1 - v^2/c^2}}, y, z\right),$$

$$G_{\zeta}(\omega, k) = \sqrt{1 - v^2/c^2} \Phi_v\left(k_x \sqrt{1 - v^2/c^2}, k_y, k_z\right) \delta(\omega - k_x v \sqrt{1 - v^2/c^2}),$$

$$\Phi_{\zeta}(k) = \sqrt{1 - v^2/c^2} \Phi_v\left(k_x \sqrt{1 - v^2/c^2}, k_y, k_z\right),$$

$$g_{\zeta}(\omega) = \frac{1}{v} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi_v\left(\frac{\omega}{v}, k_y, k_z\right) dk_y dk_z.$$

Отметим, что если покоящееся поле  $v(r)$  изотропно, то замороженное поле  $\zeta(t, r)$  статистически анизотропно.

8. Пусть  $a(r)$  и  $b(r)$  — изотропные и изотропно связанные случайные векторные поля. Найти общий вид корреляционного тензора  $\langle a_{\alpha}(r_1) b_{\beta}^*(r_2) \rangle$  в этом случае.

Решение. Для построения корреляционного тензора однородных и изотропных векторных полей мы располагаем только скалярными функциями от  $\rho \equiv |r_1 - r_2|$ , единичным вектором  $\rho/\rho$  и единичным тензором  $\delta_{\alpha\beta}$ . Поэтому общий вид корреляционного тензора будет

$$\langle a_{\alpha}(r_1) b_{\beta}^*(r_2) \rangle = F(\rho) \delta_{\alpha\beta} + [G(\rho) - F(\rho)] \frac{\rho_{\alpha} \rho_{\beta}}{\rho^2}. \quad (1)$$

Выбор скалярных коэффициентов в (1) в форме  $F$  и  $G - F$  обусловлен тем, что  $G(\rho)$  и  $F(\rho)$  описывают при этом продольную и поперечную корреляции. Действительно, рассмотрим корреляцию между  $\alpha$ -компонентами  $a_{\alpha}$  и  $b_{\alpha}$ . Если точки  $r_1$  и  $r_2$  лежат на прямой, параллельной оси  $x_{\alpha}$ , т. е.  $\rho = \rho_{\alpha}$ , то (рис. 2, а)

$$\langle a_{\alpha}(r_1) b_{\alpha}^*(r_2) \rangle = G(\rho)$$

(продольная корреляция). Если же вектор  $\rho = r_1 - r_2$  перпендикулярен оси  $x_{\alpha}$  то  $\rho_{\alpha} = 0$  (рис. 2, б) и

$$\langle a_{\alpha}(r_1) b_{\alpha}^*(r_2) \rangle = F(\rho)$$

(поперечная корреляция).

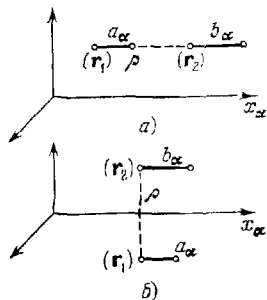


Рис. 2.

Вместо  $F(\rho)$  и  $G(\rho)$  можно ввести две другие скалярные функции, положив

$$\langle a_\alpha(r_1) b_\beta^*(r_2) \rangle = U(\rho) \delta_{\alpha\beta} + \frac{\partial^2 V(\rho)}{\partial \rho_\alpha \partial \rho_\beta}. \quad (2)$$

Сопоставление (1) и (2) показывает, что

$$F(\rho) = U(\rho) + \frac{1}{\rho} V'(\rho), \quad G(\rho) = U(\rho) + V''(\rho), \quad (3)$$

где штрих означает производную по  $\rho$ .

9. Показать, что для однородного и изотропного *потенциального* поля  $\mathbf{a}(\mathbf{r})$  ( $\text{rot } \mathbf{a} = 0$ ) продольная и поперечная функции корреляции  $F(\rho)$  и  $G(\rho)$ , введенные в предыдущей задаче, связаны соотношением (А. М. Обухов)

$$G(\rho) = [\rho F(\rho)]', \quad (1)$$

а для *соленоидального* поля ( $\text{div } \mathbf{a} = 0$ ) — соотношением (Т. Карман)

$$F(\rho) = \frac{1}{2\rho} [\rho^2 G(\rho)]'. \quad (2)$$

Решение. Пусть  $\mathbf{a} = -\nabla\varphi$  — потенциальное поле, а  $\psi(\rho)$  — корреляционная функция потенциала  $\varphi(\mathbf{r})$ . Тогда

$$\langle a_\alpha(r_1) a_\beta^*(r_2) \rangle = -\frac{\partial^2 \psi(\rho)}{\partial \rho_\alpha \partial \rho_\beta}.$$

т. е. в формуле (2) предыдущей задачи  $U=0$ ,  $V=-\psi$ . Следовательно, по формулам (3) той же задачи  $F(\rho) = \frac{1}{\rho} V'$  и  $G = V''$ , откуда и вытекает (1).

Рассматривая соленоидальное поле, умножим  $\text{div } \mathbf{a}(\mathbf{r}_1) = 0$  на  $\mathbf{a}(\mathbf{r}_2)$  и проведем статистическое усреднение, используя соотношение (2) задачи 8. Это приводит к условию  $\nabla(U + \Delta V) = 0$ , откуда с точностью до постоянной, которую всегда можно включить в  $U$ , имеем  $U + \Delta V = 0$ . Отсюда следует, что

$$U = -\Delta V = -\frac{1}{\rho^2} \frac{d}{d\rho} \left( \rho^3 \frac{dV}{d\rho} \right) = -\frac{2V'}{\rho} - V''.$$

Согласно формулам (3) задачи 8

$$F = -\frac{1}{\rho} V' - V'', \quad G = -\frac{2}{\rho} V',$$

что эквивалентно (2).

10. Показать, что при условии однородности и изотропности векторное соленоидальное поле  $\mathbf{a}(\mathbf{r})$  не коррелировано со скалярным полем  $\varphi(\mathbf{r})$  (А. М. Обухов) и с векторным потенциальным полем  $\mathbf{b}(\mathbf{r}) = -\nabla\varphi(\mathbf{r})$ .

Решение. Для однородных и изотропных полей корреляционная функция  $\langle \varphi(\mathbf{r}_1) \mathbf{a}^*(\mathbf{r}_2) \rangle$  должна быть вектором, параллельным  $\rho$ :

$$\langle \varphi(\mathbf{r}_1) \mathbf{a}^*(\mathbf{r}_2) \rangle = H(\rho) \frac{\rho}{\rho} = H(\rho),$$

где  $H(\rho)$  — скалярная функция. Поскольку для соленоидального поля  $\text{div } \mathbf{a} = 0$ , должно быть

$$\nabla_2 \langle \varphi(\mathbf{r}_1) \mathbf{a}^*(\mathbf{r}_2) \rangle = -\text{div}_2 \mathbf{B}(\rho) = 0,$$

т. е. вектор  $\mathbf{B}$  должен быть соленоидальным ( $\nabla_2$  означает дифференцирование по  $\mathbf{r}_2$ ). В отсутствие источников единственное допустимое решение есть  $\mathbf{B} = 0$ ,



что и означает некоррелированность скалярного и соленоидального векторного полей. Отсюда вытекает также некоррелированность соленоидального и потенциального векторных полей:

$$\langle b_{\beta}(\mathbf{r}_1) a_{\alpha}^*(\mathbf{r}_2) \rangle = -\frac{\partial}{\partial r_{1\beta}} \langle \varphi(\mathbf{r}_1) a_{\alpha}^*(\mathbf{r}_2) \rangle = -\frac{\partial B_{\alpha}(\rho)}{\partial r_{\beta}} = 0.$$

11. О флуктуациях с корреляционными функциями типа (2.24) и (2.25) говорят как об *одномасштабных* флуктуационных полях, имея в виду, что поведение  $\psi(\mathbf{r})$  характеризуется только одним масштабом  $l$  (пространственный спектр таких флуктуаций содержит, разумеется, множество масштабов, сосредоточенных в области  $k \lesssim 1/l$ ). Реальные функции корреляции часто имеют много характерных масштабов (при этом говорят о *многомасштабных* флуктуациях). Иногда удобно аппроксимировать такие функции суперпозицией гауссовых функций корреляции, т. е. записывать коэффициент корреляции в виде [9]

$$K(\rho) = \int_0^{\infty} f(a) e^{-\rho^2/a^2} da,$$

подчиняя, конечно, весовую функцию  $f(a)$  условию нормировки

$$K(0) = \int_0^{\infty} f(a) da = 1.$$

Обычный гауссов коэффициент корреляции  $K(\rho) = e^{-\rho^2/2l^2}$  получается отсюда при  $f(a) = \delta(a - \sqrt{2}l)$ .

Найти весовую функцию  $f(a)$ , отвечающую:

- а) экспоненциальному коэффициенту корреляции  $e^{-r/l}$ ;
- б) локально однородной и изотропной модели турбулентных флуктуаций диэлектрической проницаемости воздуха; коэффициент корреляции для таких флуктуаций в соответствии с (4.19) имеет асимптотические выражения

$$K(r) = \frac{\psi(r)}{\psi(0)} = 1 - \frac{D(r)}{D(\infty)} = \begin{cases} 1 - r^2 l_0^{-4/3} L_0^{-2/3}, & r \ll l_0, \\ 1 - (r/L_0)^{2/3}, & l_0 \ll r \ll L_0, \\ 0, & L_0 \ll r. \end{cases}$$

Ответ.

а)  $f(a) = \frac{1}{l\sqrt{\pi}} e^{-a^2/4l^2}.$

б)  $f(a) = \begin{cases} 0 & \text{при } a < l_0 \text{ и } a > L_0, \\ \frac{2}{3} a^{-1/3} L_0^{-2/3} & \text{при } l_0 \leq a \leq L_0. \end{cases}$

12. В ч. 1, § 47 было выведено выражение для функции корреляции интенсивностей двух нормальных комплексных колебаний  $\xi_1(t)$  и  $\xi_2(t)$  в предположении, что  $\bar{\xi}_1 = \bar{\xi}_2 = 0$ . Получить аналогичное выражение для случая, когда среднее значение случайного поля  $\xi(t, r)$  отлично от нуля.

Ответ. В обозначениях  $I_m = |\xi_m|^2$ , где  $\xi_m = \xi(t_m, r_m)$ ,  $m = 1, 2$ , функция корреляции интенсивностей есть

$$\psi_I(1, 2) = \langle I_1 I_2 \rangle - \bar{I}_1 \bar{I}_2 = \langle \tilde{I}_1 \tilde{I}_2 \rangle,$$

где

$$\tilde{I} = I - \bar{I} = |\tilde{\xi}|^2 = \langle |\tilde{\xi}|^2 \rangle + (\tilde{\xi} \tilde{\xi}^* + \text{к. с.}).$$

Используя результаты задачи 2 к гл. VI ч. I, получим

$$\psi_I(1, 2) = |\psi_{\zeta}(1, 2)|^2 + |\bar{\psi}_{\zeta}(1, 2)|^2 + [\psi_{\zeta}(1, 2) \bar{\zeta}_1 \bar{\zeta}_2 + \bar{\psi}_{\zeta}(1, 2) \zeta_1 \zeta_2 + \text{к. с.}].$$

При  $\bar{\zeta}_1 = \bar{\zeta}_2 = 0$  отсюда следует формула

$$\psi_I(1, 2) = |\psi_{\zeta}(1, 2)|^2 + |\bar{\psi}_{\zeta}(1, 2)|^2,$$

которая в случае  $\bar{\psi}_{\zeta} = 0$  принимает вид  $\psi_I(1, 2) = |\psi_{\zeta}(1, 2)|^2$ .

13. Пусть  $F_N[u]$  — функционал  $N$ -й степени:

$$F_N[u] = \int A(x_1, \dots, x_N) u(x_1) \dots u(x_N) dx_1 \dots dx_N.$$

Показать, что

а) значение  $F_N[u]$  зависит только от симметричной по всем аргументам части функции  $A$ ;

$$\begin{aligned} \text{б) } \frac{\delta F_N[u]}{\delta u(x)} &= \int [A(x, x_1, \dots, x_{N-1}) + A(x_1, x, x_2, \dots, x_{N-1}) + \dots \\ &\quad \dots + A(x_1, x_2, \dots, x_{N-1}, x)] u(x_1) \dots u(x_{N-1}) dx_1 \dots dx_{N-1}; \end{aligned}$$

в) в случае симметричной функции  $A$

$$\frac{\delta F_N[u]}{\delta u(x)} = N \int A(x, x_1, \dots, x_{N-1}) u(x_1) \dots u(x_{N-1}) dx_1 \dots dx_{N-1}$$

$$\left( \text{аналог формулы } \frac{dx^N}{dx} = Nx^{N-1} \right).$$

Указание. Задачу полезно решить тремя способами: исходя из определения (7.2), при помощи формулы (7.4) и при помощи правил (7.5)–(7.8).

14. Найти вариационную производную функционала действия классической механики

$$S[x(\tau)] = \int_{t_1}^{t_2} \left[ \frac{m}{2} \left( \frac{dx(\tau)}{d\tau} \right)^2 - V(x(\tau)) \right] d\tau.$$

Решение. Дифференцируя под знаком интеграла, получаем

$$\frac{\delta S}{\delta x(t)} = \int_{t_1}^{t_2} \left[ \frac{m}{2} \frac{\delta}{\delta x(t)} \left( \frac{dx(\tau)}{d\tau} \right)^2 - \frac{\delta V(x(\tau))}{\delta x(t)} \right] d\tau.$$

На основании (7.7)

$$\frac{\delta}{\delta x(t)} \left( \frac{dx(\tau)}{d\tau} \right)^2 = 2 \frac{dx(\tau)}{d\tau} \frac{\delta}{\delta x(t)} \frac{dx(\tau)}{d\tau}.$$

Так как производная  $dx(\tau)/d\tau$  есть предел отношения  $\frac{x(\tau_2) - x(\tau_1)}{\tau_2 - \tau_1}$ , можно на основании (7.5) внести операцию функционального дифференцирования под знак  $d/d\tau$ , после чего использовать (7.8):

$$\frac{\delta}{\delta x(t)} \frac{dx(\tau)}{d\tau} = \frac{\partial}{\partial \tau} \frac{\delta x(\tau)}{\delta x(t)} = \frac{\partial}{\partial \tau} \delta(\tau - t) = - \frac{\partial}{\partial t} \delta(\tau - t).$$

Для второго слагаемого в  $S[x]$ , используя (7.7) и (7.8), получаем

$$\frac{\delta V(x, (\tau))}{\delta x(t)} = V'(x(\tau)) \frac{\delta x(\tau)}{\delta x(t)} = \delta(\tau - t) V'(x(\tau)).$$

В результате

$$\begin{aligned}\frac{\delta S[x]}{\delta x(t)} &= \int_{t_1}^{t_2} \left[ -m \frac{dx(\tau)}{d\tau} \frac{\partial}{\partial t} \delta(\tau-t) - V'(x(\tau)) \delta(\tau-t) \right] d\tau = \\ &= -m \frac{\partial}{\partial t} \int_{t_1}^{t_2} \frac{dx(\tau)}{d\tau} \delta(\tau-t) d\tau - \int_{t_1}^{t_2} V'(x(\tau)) \delta(\tau-t) d\tau,\end{aligned}$$

откуда

$$\frac{\delta S[x]}{\delta x(t)} = -m \frac{d^2 x(t)}{dt^2} - V'(x(t)).$$

Принцип наименьшего действия  $\frac{\delta S[x]}{\delta x(t)} = 0$  дает уравнение движения  $m\ddot{x} = -V'(x)$ .

15. Найти  $\frac{\delta S[x]}{\delta x(t)}$  для функционала

$$S[x] = \int_{t_1}^{t_2} L(x(\tau), \dot{x}(\tau)) d\tau.$$

Ответ.

$$\frac{\delta S[x]}{\delta x(t)} = \frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}},$$

т. е. вариационная производная представляет собой левую часть уравнения Эйлера—Лагранжа.

Для получения этого результата следует использовать формулы (7.5)—(7.8), подобно тому как это сделано в предыдущей задаче.

16. Вывести формулу для  $\psi_3(x_1, x_2, x_3)$ , аналогичную формуле (7.17) для  $\psi_2(x_1, x_2)$ .

Ответ.

$$\begin{aligned}\psi_3(x_1, x_2, x_3) &= \langle \xi(x_1) \xi(x_2) \xi(x_3) \rangle - \bar{\xi}(x_1) \langle \xi(x_2) \xi(x_3) \rangle - \bar{\xi}(x_2) \langle \xi(x_1) \xi(x_3) \rangle - \\ &\quad - \bar{\xi}(x_3) \langle \xi(x_1) \xi(x_2) \rangle + 2\bar{\xi}(x_1) \bar{\xi}(x_2) \bar{\xi}(x_3).\end{aligned}$$

17. Исходя из выражений (7.20) для центральных моментов гауссова случайного поля  $\xi(x)$ , получить характеристический функционал для флуктуаций  $\xi(x)$ .

Решение. Согласно (7.14) имеем

$$\Phi_{\xi}^{-}[u] = 1 + \sum_{k=1}^{\infty} \frac{i^k}{k!} \int \langle \bar{\xi}(x_1) \dots \bar{\xi}(x_k) \rangle u(x_1) \dots u(x_k) dx_1 \dots dx_k,$$

причем, в силу (7.20), в этой сумме присутствуют только члены с  $k=2m$ :

$$\Phi_{\xi}^{-}[u] = 1 + \sum_{m=1}^{\infty} \frac{(-1)^m}{(2m)!} \int \langle \bar{\xi}(x_1) \dots \bar{\xi}(x_{2m}) \rangle u(x_1) \dots u(x_{2m}) dx_1 \dots dx_{2m}.$$

Используя теперь вторую формулу (7.20), получаем

$$\begin{aligned}\Phi_{\xi}^{-}[u] &= 1 + \sum_{m=1}^{\infty} \frac{(-1)^m}{(2m)!} \sum_{p, n} \int \langle \bar{\xi}(x_{\alpha_1}) \bar{\xi}(x_{\alpha_2}) \dots \bar{\xi}(x_{\alpha_{2m-1}}) \bar{\xi}(x_{\alpha_{2m}}) \rangle u(x_{\alpha_1}) u(x_{\alpha_2}) \dots \\ &\quad \dots u(x_{\alpha_{2m-1}}) u(x_{\alpha_{2m}}) dx_1 \dots dx_{2m}\end{aligned}$$

Интеграл кратности  $(2m)$  распадается на произведение  $m$  двукратных интегралов, отличающихся друг от друга лишь обозначением переменных интегрирования. Каждый из двукратных интегралов имеет вид

$$I = \int \int \langle \tilde{\xi}(x') \tilde{\xi}(x'') \rangle u(x') u(x'') dx' dx''.$$

Следовательно, каждый из членов внутренней суммы равен  $I^m$ , а так как сумма содержит всего  $(2m-1)!!$  членов, функционал  $\Phi_{\tilde{\xi}}[u]$  равен

$$\Phi_{\tilde{\xi}}[u] = 1 + \sum_{m=1}^{\infty} \frac{(-1)^m}{(2m)!} (2m-1)!! I^m,$$

или, поскольку  $(2m)! = (2m)!! \cdot (2m-1)!! = 2^m m! (2m-1)!!$ ,

$$\Phi_{\tilde{\xi}}[u] = 1 + \sum_{m=1}^{\infty} \frac{(-1)^m}{m!} \left( \frac{I}{2} \right)^m = \exp \left\{ -\frac{1}{2} I \right\}.$$

Таким образом,

$$\Phi_{\tilde{\xi}}[u] = \exp \left\{ -1/2 \int \int \psi_{\tilde{\xi}}(x_1, x_2) u(x_1) u(x_2) dx_1 dx_2 \right\}.$$

Очевидно, эта формула является частным случаем (7.18), отвечающим нулевому среднему значению, как это и должно было получиться для  $\tilde{\xi} = \xi - \bar{\xi}$ , поскольку  $\langle \tilde{\xi} \rangle = 0$ . Тем самым доказана эквивалентность определений гауссова случайного поля при помощи равенств (7.20) и при помощи характеристического функционала (7.18).

18. Найти характеристический функционал пуассоновского случайного поля (7.21).

Решение. По определению пуассоновского поля

$$\Phi[u] = \left\langle \left\langle \exp \left\{ i \int dx u(x) \sum_{k=0}^m A_k g(x-x_k) \right\} \right\rangle_{A_k} \right\rangle_{x_k},$$

где каждая из угловых скобок с индексом внизу означает усреднение по соответствующим случайным величинам. Очевидно,

$$\Phi[u] = \left\langle \left\langle \prod_{k=0}^m \exp \left\{ i A_k \int g(x-x_k) u(x) dx \right\} \right\rangle_{A_k} \right\rangle_{x_k}.$$

Выполним усреднение по  $A_k$ . Так как  $\langle \exp \{ i A_k \lambda \} \rangle_{A_k} = \chi(\lambda)$ , где  $\chi(\lambda)$  — характеристическая функция для  $A_k$ , находим

$$\Phi[u] = \left\langle \prod_{k=0}^m \chi \left( \int g(x-x_k) u(x) dx \right) \right\rangle_{x_k}.$$

Далее произведем усреднение по  $x_k$ . Оно сводится к интегрированию каждого сомножителя по  $x_k$  с весом  $V^{-1}$  по области  $V$ :

$$\begin{aligned} \Phi[u] &= \left\langle V^{-1} \int_V \chi \left( \int g(x-x_1) u(x) dx \right) dx_1 \dots \right. \\ &\quad \left. \dots V^{-1} \int_V \chi \left( \int g(x-x_m) u(x) dx \right) dx_m \right\rangle_m = \\ &= \left\langle \left[ V^{-1} \int_V \chi \left( \int g(x-x') u(x) dx \right) dx' \right]^m \right\rangle_m. \end{aligned}$$

Обозначим здесь выражение в квадратных скобках через  $a$ . Последнее усреднение (по распределению Пуассона для  $m$ ) дает

$$\langle a^m \rangle_m = \sum_{m=0}^{\infty} P(m) a^m = e^{-\bar{m}} \sum_{m=0}^{\infty} \frac{(\bar{m})^m}{m!} a^m = e^{-\bar{m}} e^{a\bar{m}} = \exp \{ \bar{m} (a-1) \}.$$

Таким образом, получаем

$$\Phi[u] = \exp \left\{ \bar{m} \left[ V^{-1} \int_V \chi \left( \int g(x-x') u(x) dx \right) dx' - 1 \right] \right\},$$

что, в силу тождества

$$1 = V^{-1} \int_V 1 \cdot dx',$$

можно записать в виде

$$\Phi[u] = \exp \left\{ \frac{\bar{m}}{V} \int_V \left[ \chi \left( \int g(x-x') u(x) dx \right) - 1 \right] dx' \right\}.$$

Вводя  $\nu = \bar{m}/V$  — среднее число случайных точек в единице объема, приходим к формуле (7.22).

19. Исходя из формулы (7.22), получить кумулянтные функции пуассоновского случайного поля.

Решение. Имеем

$$\Theta[u] = \ln \Phi[u] = \nu \int \left[ \chi \left( \int g(x-x') u(x) dx \right) - 1 \right] dx'.$$

Подстановка разложения характеристической функции

$$\chi(\lambda) = \sum_{k=0}^{\infty} \frac{(i\lambda)^k}{k!} \langle A^k \rangle$$

даст

$$\Theta[u] = \nu \sum_{k=1}^{\infty} \frac{i^k}{k!} \langle A^k \rangle \int \left[ \int g(x-x') u(x) dx \right]^k dx'.$$

Запишем  $k$ -ю степень интеграла в виде  $k$ -кратного интеграла и изменим порядок интегрирования по  $x'$  и  $x_l$  ( $l=1, 2, \dots, k$ ):

$$\begin{aligned} \Theta[u] &= \nu \sum_{k=1}^{\infty} \frac{i^k}{k!} \langle A^k \rangle \int dx' \int dx_1 \dots \int dx_k g(x_1-x') \dots g(x_k-x') u(x_1) \dots u(x_k) = \\ &= \sum_{k=1}^{\infty} \frac{i^k}{k!} \int u(x_1) \dots u(x_k) dx_1 \dots dx_k \left\{ \nu \langle A^k \rangle \int dx' g(x_1-x') \dots g(x_k-x') \right\} \end{aligned}$$

Сравнивая это выражение с (7.15), получаем, что

$$\psi_k(x_1, \dots, x_k) = \nu \langle A^k \rangle \int g(x_1-x') \dots g(x_k-x') dx'.$$

20. Показать, что если  $A$  имеет гауссово распределение вероятностей и  $\bar{A} = 0$ , то при  $v \rightarrow \infty$ ,  $\bar{A}^2 \rightarrow 0$  и  $v\bar{A}^2 = \text{const}$  пуассоновское случайное поле стремится к гауссову. Показать, что аналогичный результат можно получить, если считать, что  $A$  имеет плотность вероятностей  $w(A) = (2A_0)^{-1} \times \chi \exp \{-|A|/A_0\}$ , т. е. нормальность распределения вероятностей для амплитуд несущественна.

Указание. Воспользоваться формулой (7.23) для кумулянтов пуассоновского случайного поля.

21. Исходя из (7.27), получить формулу для вычисления среднего значения от произведения пуассоновского случайного поля  $\xi(x)$  на функционал  $R[\xi]$ .

Решение. Согласно (7.22)

$$\Theta[u] = v \int \left[ \chi \left( \int g(x'' - x') u(x'') dx'' \right) - 1 \right] dx'.$$

Действуя на это равенство оператором  $\frac{1}{i} \frac{\delta}{\delta u(x)}$ , найдем

$$\Omega[x, u] = \frac{v}{i} \int dx' \chi' \left( \int g(x'' - x') u(x'') dx'' \right) g(x - x').$$

Но

$$\chi(\lambda) = \int e^{i\lambda A} w(A) dA, \quad \frac{1}{i} \chi'(\lambda) = \int e^{i\lambda A} w(A) A dA.$$

Полагая  $\lambda = \frac{1}{i} \int dx'' g(x'' - x') \frac{\delta}{\delta \xi(x')}$ , получаем

$$\Omega \left[ x, \frac{1}{i} \frac{\delta}{\delta \xi} \right] = v \int dx' g(x - x') \int w(A) A \exp \left\{ A \int dx'' g(x'' - x') \frac{\delta}{\delta \xi(x'')} \right\} dA.$$

В результате формула (7.27), с учетом операторной записи функционального ряда Тейлора (7.10), принимает вид

$$\langle \xi(x) R[\xi] \rangle = v \int g(x - x') dx' \int w(A) A \langle R[\xi(\bar{x}) + A g(\bar{x} - x')] \rangle dA.$$

22. Пусть  $X_i(x, t)$  — детерминированные функции,  $f_i(x, t)$  — случайные гауссовы функции  $(n+1)$  переменных  $(x, t)$ , причем

$$\langle f_i \rangle = 0, \quad \langle f_i(x, t) f_k(x', t') \rangle = 2F_{ik}(x, x', t) \delta(t - t'). \quad (1)$$

Пусть функции  $\xi_i(t)$  подчинены системе динамических уравнений

$$\frac{d\xi_i(t)}{dt} = X_i(\xi(t), t) + f_i(\xi(t), t) \quad (2)$$

с начальным условием  $\xi_i(0) = \xi_i^0$ .

Вывести дифференциальное уравнение для плотности вероятностей  $w(x, t) = \langle \delta(\xi(t) - x) \rangle$  решения уравнений (2).

Решение. Продифференцируем  $w(x, t)$  по  $t$ :

$$\frac{\partial w(x, t)}{\partial t} = \left\langle \frac{\partial}{\partial t} \delta(\xi(t) - x) \right\rangle = \left\langle - \frac{\partial \delta(\xi(t) - x)}{\partial x_i} \frac{d\xi_i(t)}{dt} \right\rangle,$$

и подставим сюда  $d\xi_i(t)/dt$  из уравнений (2):

$$\begin{aligned} \frac{\partial w(x, t)}{\partial t} &= - \left\langle \frac{\partial \delta(\xi(t) - x)}{\partial x_i} [X_i(\xi(t), t) + f_i(\xi(t), t)] \right\rangle = \\ &= - \frac{\partial}{\partial x_i} \langle \delta(\xi(t) - x) [X_i(\xi(t), t) + f_i(\xi(t), t)] \rangle. \end{aligned}$$

Последнее выражение — результат вынесения  $\frac{\partial}{\partial x_i}$  [за скобку и за знак среднего, поскольку единственный сомножитель, зависящий от  $x$ , — это дельта-функция. Используя равенство

$$\delta(\xi(t) - x) F(\xi(t)) = \delta(\xi(t) - x) F(x),$$

получаем

$$\begin{aligned} \frac{\partial \omega(x, t)}{\partial t} &= - \frac{\partial}{\partial x_i} \langle \delta(\xi(t) - x) [X_i(x, t) + f_i(x, t)] \rangle = \\ &= - \frac{\partial}{\partial x_i} [X_i(x, t) \langle \delta(\xi(t) - x) \rangle + \langle f_i(x, t) \delta(\xi(t) - x) \rangle], \end{aligned}$$

где мы вынесли за знак среднего неслучайный множитель  $X_i(x, t)$ . Так как  $\langle \delta(\xi(t) - x) \rangle = \omega(x, t)$ , получаем уравнение

$$\frac{\partial \omega(x, t)}{\partial t} + \frac{\partial}{\partial x_i} [X_i(x, t) \omega(x, t)] = - \frac{\partial Q_i}{\partial x_i}, \quad (3)$$

где

$$Q_i(x, t) = \langle f_i(x, t) \delta(\xi(t) - x) \rangle. \quad (4)$$

Уравнение (3) не является замкнутым, так как кроме искомой функции  $\omega$  в него входят еще неизвестные функции  $Q_i$ . Дальнейшая задача состоит в том, чтобы выразить  $Q_i$  через  $\omega$ . Используем для этого формулу Фуруцу—Новикова (7.33), которая в нашем случае и с учетом (1) имеет вид

$$\begin{aligned} \langle f_i(x, t) R[f] \rangle &= \langle f_i(x, t) \rangle \langle R \rangle + \\ &+ \int dx' \int dt' \langle \tilde{f}_i(x, t) \tilde{f}_k(x', t') \rangle \left\langle \frac{\delta R[f]}{\delta f_k(x', t')} \right\rangle = \\ &= 2 \int dx' \int dt' F_{ik}(x, x', t) \delta(t - t') \left\langle \frac{\delta R[f]}{\delta f_k(x', t')} \right\rangle. \end{aligned} \quad (5)$$

Для того чтобы получить  $Q_i(x, t)$ , надо, согласно (4), положить  $R[f] = \delta(\xi(t) - x)$ . Это вполне законное равенство, поскольку  $\xi(t)$ , будучи решением системы уравнений (2), является функционалом от  $f_i$ .

Чтобы найти  $\langle \delta R / \delta f_k \rangle$ , обратимся снова к уравнениям (2). Проинтегрировав их по  $t$ , находим

$$\xi_i(t) = \xi_i^0 + \int_0^t [X_i(\xi(\tau), \tau) + f_i(\xi(\tau), \tau)] d\tau. \quad (6)$$

Отсюда видно, что  $\xi_i(t)$  зависит от значений  $f_k(\tau)$ , принимаемых только на интервале  $0 \leq \tau \leq t$ . Поэтому

$$\frac{\delta \xi_i(t)}{\delta f_k(t')} = 0, \text{ если } t' < 0 \text{ или } t' > t. \quad (7)$$

Формула (7) показывает, что решения системы уравнений (2) удовлетворяют принципу динамической причинности. Запишем (6) в виде

$$\xi_i(t) = \xi_i^0 + \int_0^t d\tau \int dx \delta(\xi(\tau) - x) [X_i(x, \tau) + f_i(x, \tau)]$$

и применим к этим выражениям оператор  $\delta/\delta f_j(x', t')$ , где  $0 < t' < t$ :

$$\frac{\delta \xi_i(t)}{\delta f_j(x', t')} = \int_0^t d\tau \int dx \left\{ -\frac{\partial \delta(\xi(\tau) - x)}{\partial x_k} \frac{\delta \xi_k(\tau)}{\delta f_j(x', t')} [X_i(x, \tau) + f_i(x, \tau)] + \right. \\ \left. + \delta(\xi(\tau) - x) \frac{\delta f_i(x, \tau)}{\delta f_j(x', t')} \right\}.$$

Первое слагаемое в фигурной скобке проинтегрируем по частям. Во втором же слагаемом учтем, что, согласно (7.32),

$$\frac{\partial f_i(x, \tau)}{\delta f_j(x', t')} = \delta_{ij} \delta(x - x') \delta(\tau - t'),$$

и используем при интегрировании наличие этих дельта-функций. В результате

$$\frac{\delta \xi_i(t)}{\delta f_j(x', t')} = \delta_{ij} \delta(\xi(t') - x') + \\ + \int_0^t d\tau \int dx \delta(\xi(\tau) - x) \frac{\delta \xi_k(\tau)}{\delta f_j(x', t')} \frac{\partial}{\partial x_k} [X_i(x, \tau) + f_i(x, \tau)].$$

Но, в силу условия причинности (7),  $\frac{\delta \xi_k(\tau)}{\delta f_j(x', t')} = 0$  при  $\tau < t'$ , что позволяет заменить нижний предел интеграла по  $\tau$  на  $t'$ :

$$\frac{\delta \xi_i(t)}{\delta f_j(x', t')} = \delta_{ij} \delta(\xi(t') - x') + \int_{t'}^t d\tau \dots$$

Тогда переход к пределу при  $t' \rightarrow t$  обращает интеграл в нуль и мы получаем точную формулу

$$\frac{\delta \xi_i(t)}{\delta f_j(x', t)} = \delta_{ij} \delta(\xi(t) - x'), \quad (8)$$

являющуюся следствием динамических уравнений (2).

Вернемся к формуле (5) для  $Q_i$ . Так как при  $t' > t$  в силу (7) имеем  $\langle \frac{\delta R}{\delta f_k} \rangle = 0$ , интегрирование по  $t'$  в (5) происходит в пределах  $(0, t)$ . При выполнении интегрирования по  $t'$  следует учесть, что  $\int_0^t \delta(t - t') dt' = 1/2$ , поскольку фигурирующая здесь дельта-функция предполагается пределом четной корреляционной функции. Поэтому

$$Q_i(x, t) = \int F_{ik}(x, x', t) \left\langle \frac{\delta R[f]}{\delta f_k(x', t)} \right\rangle dx'. \quad (9)$$

Так как  $R[f] = \delta(\xi(t) - x)$ , имеем

$$\frac{\delta R[f]}{\delta f_k(x', t)} = \frac{\delta}{\delta f_k(x', t)} \delta(\xi(t) - x) = -\frac{\partial \delta(\xi(t) - x)}{\partial x_k} \frac{\delta \xi_k(t)}{\delta f_k(x', t)},$$



или, в силу (8),

$$\begin{aligned}\frac{\delta R[f]}{\delta f_k(x', t)} &= -\delta_{ik} \delta(\xi(t) - x') \frac{\partial \delta(\xi(t) - x)}{\partial x_i} = \\ &= -\frac{\partial}{\partial x_k} [\delta(\xi(t) - x') \delta(\xi(t) - x)] = -\frac{\partial}{\partial x_k} [\delta(x - x') \delta(\xi(t) - x)].\end{aligned}$$

Усредняя это равенство, получаем формулу

$$\left\langle \frac{\delta R[f]}{\delta f_k(x', t)} \right\rangle = -\frac{\partial}{\partial x_k} [\delta(x - x') w(x, t)],$$

с учетом которой выражение (9) принимает вид

$$Q_i(x, t) = -\int F_{ik}(x, x', t) \frac{\partial}{\partial x_k} [w(x, t) \delta(x - x')] dx'.$$

Записывая интегрант в виде

$$\begin{aligned}-F_{ik}(x, x', t) \frac{\partial}{\partial x_k} [w(x, t) \delta(x - x')] &= \\ &= -\frac{\partial}{\partial x_k} [F_{ik}(x, x', t) w(x, t) \delta(x - x')] + \frac{\partial F_{ik}(x, x', t)}{\partial x_k} w(x, t) \delta(x - x'),\end{aligned}$$

получаем

$$\begin{aligned}Q_i(x, t) &= -\frac{\partial}{\partial x_k} \int F_{ik}(x, x', t) w(x, t) \delta(x - x') dx' + \\ &+ \int \frac{\partial F_{ik}(x, x', t)}{\partial x_k} w(x, t) \delta(x - x') dx' = \\ &= -\frac{\partial}{\partial x_k} [F_{ik}(x, x, t) w(x, t)] + \left[ \frac{\partial F_{ik}(x, x', t)}{\partial x_k} \right]_{x'=x} w(x, t).\end{aligned}$$

Таким образом, функция  $Q_i(x, t)$  выражена через искомую плотность вероятностей  $w(x, t)$ , после чего (3) становится замкнутым уравнением:

$$\frac{\partial w(x, t)}{\partial t} + \frac{\partial A_i(x, t) w}{\partial x_i} = \frac{\partial^2 F_{ik}(x, x, t) w}{\partial x_i \partial x_k}, \quad (10)$$

где введено обозначение

$$A_i(x, t) = X_i(x, t) + \left[ \frac{\partial F_{ik}(x, x', t)}{\partial x_k} \right]_{x'=x}.$$

Уравнение (10) представляет собой уравнение Эйнштейна—Фоккера, соответствующее динамической системе (2). В ч. 1, § 36 оно было приведено без вывода, причем множители при  $\delta(t-t')$  в корреляционной матрице воздействий  $f_i(x, t)$  были обозначены через  $C_{ik}(x, x', t) = 2F_{ik}(x, x', t)$ .

## Глава II

### ИЗЛУЧЕНИЕ И ДИФРАКЦИЯ СЛУЧАЙНЫХ ВОЛНОВЫХ ПОЛЕЙ

#### § 8. Основные типы статистических волновых задач

Среди разнообразных случайных полей, с которыми имеет дело статистическая радиофизика, *волновые* поля занимают центральное место. Мы тоже сосредоточим внимание на волновых (в первую очередь электромагнитных) полях и ограничимся при этом только *линейными* и *неквантовыми* задачами<sup>1)</sup>. Весьма широкий класс таких задач можно сформулировать следующим образом.

Пусть распространение волн той или иной физической природы (электромагнитных, упругих, поверхностных и т. д.) описывается линейным пространственно-временным оператором  $\hat{L}$  (обычно дифференциальным, реже — интегро-дифференциальным), так что волновое поле  $u_a$  удовлетворяет уравнению

$$\hat{L}u = q, \quad (8.1)$$

где функция  $q(t, r)$  описывает источники волн. Поля  $u$  и  $q$  могут быть и многокомпонентными (в частности, векторными), и тогда  $\hat{L}$  — операторная матрица (в частности, тензор). Во многих задачах пространственная область, в которой рассматривается поле  $u$ , выделенная некоторой поверхностью  $S_0$ , не содержит источ-

<sup>1)</sup> Укажем некоторые обзоры и монографии по нелинейным волновым задачам и по квантовой статистике излучения. Статистическим явлениям в *нелинейной оптике* посвящены книга [1] и обзор [2]. Обширные исследования ведутся по теории *слабой турбулентности* — явлению, которое происходит во многих ситуациях, в том числе в случае волн в плазме [3, 4], на поверхности жидкости [4, 5], в активных распределенных системах и т. д. Статистические вопросы *нелинейной акустики* рассмотрены в книге [45]. Вопросы *квантовых флуктуаций* электромагнитного поля, ставшие особенно актуальными в связи с развитием лазерной техники, освещены в работах [6—10].

ников ( $q = 0$ ), а задано *первичное* волновое поле  $u_0$ , приходящее в эту область извне. Тогда уравнение (8.1) однородно:

$$\hat{L}u = 0, \quad (8.2)$$

но на  $S_0$  заданы значения *первичного* поля (или его производных), например:

$$v \equiv u_0(t, \mathbf{r})|_{\mathbf{r} \in S_0}. \quad (8.3)$$

(Обычно в этом случае говорят, что на  $S_0$  заданы «виртуальные» источники поля.) Искомым является здесь рассеянное или дифракционное поле, т. е. это задачи теории дифракции.

При наличии внутри поверхности  $S_0$  границ раздела между разными средами или телами (поверхности раздела  $S$ ) поле  $u$  должно удовлетворять еще определенным *граничным условиям*. Если поверхность  $S_0$  не замкнута или же охватывает все пространство, так что волны от реальных источников, расположенных в конечной области, могут уходить в бесконечность, то должны также выполняться известные *условия излучения* (на достаточно больших расстояниях от источников должны существовать только убегающие волны).

В задачах прикладного характера часто представляет интерес *измерение* излученного или дифракционного поля — для получения информации об источниках поля, о рассеивающих телах или о среде, в которой распространяются волны. Тогда в описанию схемы может быть включен еще *приемник* излучения<sup>1)</sup>, а также разного рода *помехи* как внешнего (по отношению к приемнику), так и внутреннего происхождения. Отклик приемника  $w$  будет зависеть и от измеряемого поля  $u$ , и от помех  $\xi$ :

$$w = \hat{w}(u, \xi), \quad (8.4)$$

где  $\hat{w}$  — в общем случае нелинейный оператор.

Все, что было сказано в ч. 1, § 33 об обыкновенных стохастических дифференциальных уравнениях, теперь, когда мы рассматриваем случайные поля, переносится на уравнения в частных производных. Статистические волновые задачи ставятся теми же уравнениями и условиями, что и динамические, но теперь это будут *стохастические* уравнения и условия, т. е. уравнения и условия для отдельных реализаций случайного поля  $u$ . Другими словами, фигурирующие в задаче параметры, функции и операторы теперь случайны (все или их часть) и, соответственно, заданы своими распределениями вероятностей.

<sup>1)</sup> Под приемником подразумевается все измерительное устройство в целом, например радиантенна вместе с усилительным трактом и регистрирующим устройством или фотодетектор с измерительными приборами и т. п.

Поэтому гораздо большее разнообразие возможностей для пространственно-временных полей (по сравнению с процессами во времени) в равной мере затрагивает как динамические, так и статистические задачи.

В соответствии с описанной постановкой динамической волновой задачи случайными могут быть

1) *источники поля* (реальные или виртуальные, так что можно различать заданную «статистику источников»  $q$  и «статистику первичного поля»  $v$ );

2) *свойства среды* (задана «статистика среды», а значит, оператора  $\hat{L}$ );

3) *форма и положение границ раздела  $S$*  (задана «статистика границ»);

4) *условия приема и регистрации волн* (заданы «статистика приемника» — оператора  $\hat{w}$  и «статистика помех»  $\xi$ ).

К этим четырем основным статистическим схемам, которые мы назовем первичными, фактически сводится постановка подавляющего большинства задач статистической волновой теории. Конечно, возможны задачи и смешанного типа, например о тепловом излучении в случайно-неоднородной среде [11], но пока таких задач рассмотрено немного (краткий их обзор приведен в [12]).

Если бы мы располагали *точным решением динамической задачи*, например некоторым интегральным представлением искомого поля  $u$  в виде

$$u = \hat{G}q, \quad (8.5)$$

где  $\hat{G}q = \hat{L}^{-1}q$  — решение неоднородного уравнения (8.1), записанное, скажем, через функцию Грина, учитывающую все граничные и другие условия, то вычисление моментов поля свелось бы к усреднению произведений вида  $u(1)u(2) \dots u(n)$  по совместному распределению всех фигурирующих в задаче случайных параметров и функций, характеризующих статистику источников, среды, границ раздела и т. д. Но в реальных ситуациях этот идеал осуществляется не часто, например, в задачах о возбуждении полей случайными источниками, которые рассматриваются в данной главе. Чаще же всего мы не умеем находить точное решение при любых детерминированных функциях  $u$ , тем самым, при произвольных реализациях всех случайных величин и функций, в силу чего приходится уже на этапе решения динамической задачи обращаться к разного рода *приближенным* методам. Эти разнообразные методы и приемы приурочены к конкретным особенностям задачи.

Флуктуации случайных параметров и функций могут быть (в каких-то характерных масштабах) большими и малыми, плав-

ными, медленными и, наоборот, резкими, быстрыми; корреляция может быть сильной, «далекой», или же слабой, «короткой», и т. п. Эти различия требуют использования разных приближенных подходов и приводят к многочисленным вторичным статистическим схемам, связанным уже с теми или иными приближенными методами решения.

Следует заметить, что во многих случаях, в особенности когда нас интересуют только моменты поля, этап отыскания динамического решения (для последующего вычисления моментов) опускается и речь идет о выводе уравнений и условий для самих моментов (исходя из уравнений и условий для поля  $u$ ). Однако и в такой постановке практически нельзя обойтись без максимально возможного упрощения исходных уравнений для  $u$ , заранее учитывающего особенности флуктуаций и характер детерминированных функций.

Выделенные выше четыре первичные статистические схемы отражают лишь фактически часто встречающееся разделение параметров и функций, входящих в условия задачи, на детерминированные и случайные. Остановимся коротко на этих схемах и укажем некоторые примеры относящихся к ним задач.

В схеме 1), если присутствуют реальные источники, мы имеем дело с *неоднородным* уравнением (8.1), в котором статистически задана *правая часть*  $q$ . *Однородные* граничные условия детерминированы. Задач такого типа много и в радиофизике, и в оптике, и в акустике. Они охватывают, в частности, статистическую теорию антенн и теорию тепловых флуктуаций в распределенных системах.

С виртуальными случайными источниками мы сталкиваемся, очевидно, во всех задачах о дифракции случайных (иначе — частично когерентных) полей<sup>1)</sup>, когда однородное уравнение (8.2) и все необходимые условия детерминированы, за исключением случайного первичного поля  $u_0$ . Такие задачи типичны прежде всего для оптики (формирование оптического и голографического изображений, действие интерферометров и др.), но с ними приходится иметь дело и в радиодиапазоне (в частности, в радиоастрономии), и при дифракции рентгеновских волн.

Задачи типа 1) мы рассмотрим в данной главе, но тепловые флуктуации в распределенных системах, ввиду важности и специфичности этого круга вопросов, мы выделяем в самостоятельную (следующую) главу.

Схема 2) охватывает проблему распространения и дифракции волн в *случайно-неоднородных* средах (случайный оператор  $\hat{L}$ ). Эти

---

<sup>1)</sup> Применительно к волновым полям термины «случайный» и «частично когерентный» равнозначны.

вопросы представляют большой интерес для радиосвязи, лазерной связи, гидроакустики, радиоастрономии, диагностики плазмы и т. п., и им уделена поэтому половина данной части книги.

К схеме 3) относятся волновые задачи при наличии тел имеющих *случайную форму* или занимающих *случайное положение*. Речь может идти, в частности, о граничных поверхностях со множеством случайных неровностей (так называемые *шероховатые* или *статистически неровные* поверхности). Различные методы расчета рассеяния волн на таких поверхностях рассмотрены в гл. IX.

В задачах о телах, занимающих случайное положение в пространстве, речь может идти о рассеянии как на одном или немногих телах, так и на очень большой совокупности дискретных вкраплений (осадки, туман, аэрозоли и т. п.). Последняя весьма общая задача требует, вообще говоря, учета многократного рассеяния. Мы ограничимся ее рассмотрением лишь в приближении однократного рассеяния (§ 31).

Наконец, схема 4) охватывает многочисленные задачи приемы и обработки информации о волновых полях при наличии помех. Если статистические свойства поля  $u$ , помех  $\xi$  и оператора  $\hat{w}$ , описывающего приемник, известны, то, в соответствии с (8.4), в принципе можно рассчитать статистические характеристики отклика приемника  $w$ .

Однако более важным и вместе с тем более сложным является другой вопрос — о выборе оптимального (в каком-то определенном смысле) способа приема при наличии помех, т. е. вопрос о нахождении *оптимального* оператора  $\hat{w}$ . Примером задачи такого типа может служить проблема восстановления формы объекта по его изображению, представляющая первостепенный интерес для оптики, радиоастрономии, радиолокации, гидроакустики и т. д. Эта проблема оптимального приема случайных полей требует привлечения идей и методов теории информации — возможно, даже в большей степени, чем вопросы обработки случайных процессов. Но, как и в ч. I, мы не будем углубляться в эти проблемы, поскольку они ближе по своему характеру к «радиоматематике», а не к радиофизике. Ряд вопросов оптимального приема и пространственной фильтрации рассматривается в работах [13—20].

## § 9. Случайные волны в неограниченной однородной среде

Задачи возбуждения полей случайными источниками (реальными или виртуальными), отнесенные нами к статистической схеме I), принадлежат к тем весьма немногочисленным проблемам статистической волновой теории, которые допускают, по

существо, универсальный подход. Это обусловлено тем, что поле  $u$  связано с источниками *линейным детерминированным* оператором, который, в принципе, может быть обращен.

При возбуждении поля реальными источниками в силу (8.5) любые моменты поля могут быть получены усреднением произведений вида  $u(1)u(2) \dots u(n)$  лишь по ансамблю случайных источников  $q$ . В частности, для двух низших моментов имеем

$$\langle u \rangle = \hat{G} \langle q \rangle, \quad \langle u_1 u_2^* \rangle = \hat{G}_1 \hat{G}_2^* \langle q_1 q_2^* \rangle. \quad (9.1)$$

В случае *виртуальных* источников, когда в соответствии с (8.3) заданы значения  $v$  первичного поля (или его производных) на некоторой поверхности  $S_0$ , а искомое поле выражается через  $v$  при помощи *линейного детерминированного* оператора  $\hat{\mathcal{P}}$ :

$$u = \hat{\mathcal{P}}v, \quad (9.2)$$

искомые моменты  $u$  связаны с известными моментами  $v$  линейными соотношениями, подобными (9.1):

$$\langle u \rangle = \hat{\mathcal{P}} \langle v \rangle, \quad \langle u_1 u_2^* \rangle = \hat{\mathcal{P}}_1 \hat{\mathcal{P}}_2^* \langle v_1 v_2^* \rangle, \dots \quad (9.3)$$

Соотношения вида (9.1) или (9.3), в принципе, дают полное решение статистических задач схемы 1), поскольку совокупность всех статистических моментов однозначно определяет всю совокупность  $n$ -мерных плотностей вероятностей случайного поля. Однако эта формально простая процедура фактически реализуема крайне редко. Удобные для физического анализа выражения для моментов поля  $u$  в большинстве случаев удается получать только для низших моментов и лишь при использовании тех или иных приближений для обратных операторов  $\hat{G}$  или  $\hat{\mathcal{P}}$ . Получение же плотности вероятностей поля осуществимо обычно лишь при условиях, когда применима центральная предельная теорема.

С необходимостью прибегать к различным приближениям мы сталкиваемся даже в простейшем случае *скалярного* волнового поля в *однородной безграничной* среде, когда мы располагаем сравнительно простыми *точными* выражениями для операторов  $\hat{G}$  и  $\hat{\mathcal{P}}$ . Приведем относящиеся к этому случаю и необходимые для дальнейшего динамические соотношения и выясним на примере скалярного поля ряд свойств случайных волновых полей. В последующих параграфах мы обратимся к конкретным радиофизическим задачам, относящимся к схеме 1).

В однородной и стационарной среде без дисперсии и поглощения скалярное поле  $u$  удовлетворяет волновому уравнению

$$\hat{L}u = \left( \Delta - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) u(t, r) = q(t, r), \quad (9.4)$$

где  $c$  — скорость распространения волн, а  $q$  — поле случайных источников. Решение уравнения (9.4), удовлетворяющее условиям излучения на бесконечности, имеет, как известно, вид

$$u(t, \mathbf{r}) = -\frac{1}{4\pi} \int \frac{q(t - R/c, \mathbf{r}')}{R} d^3r', \quad (9.5)$$

где  $R = |\mathbf{r} - \mathbf{r}'|$ . Это частный случай линейной связи (8.5) поля с источниками.

Во многих задачах удобнее рассматривать не само случайное поле  $u(t, \mathbf{r})$ , а его спектральную амплитуду

$$u(\omega, \mathbf{r}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} u(t, \mathbf{r}) e^{i\omega t} dt, \quad (9.6)$$

которая, в силу (9.4), удовлетворяет уравнению

$$\Delta u(\omega, \mathbf{r}) + k^2 u(\omega, \mathbf{r}) = q(\omega, \mathbf{r}), \quad k = \omega/c, \quad (9.7)$$

где  $q(\omega, \mathbf{r})$  — спектральная амплитуда  $q(t, \mathbf{r})$ . Для спектральных амплитуд решение (9.5) принимает вид

$$u(\omega, \mathbf{r}) = -\frac{1}{4\pi} \int \frac{q(\omega, \mathbf{r}') e^{ikR}}{R} d^3r' \equiv \int G(\mathbf{r}, \mathbf{r}') q(\omega, \mathbf{r}') d^3r'. \quad (9.8)$$

Здесь введена функция Грина для неограниченной однородной среды

$$G(\mathbf{r}, \mathbf{r}') = -\frac{1}{4\pi} \frac{e^{ikR}}{R}, \quad R = |\mathbf{r} - \mathbf{r}'|, \quad (9.9)$$

удовлетворяющая уравнению

$$\Delta G + k^2 G = \delta(\mathbf{r} - \mathbf{r}').$$

При помощи (9.8) легко записать выражения для всех моментов поля.

Если источники сосредоточены в ограниченной области, скажем в пределах сферы радиуса  $a$ , а нас интересует поле  $u$  в дальней (фраунгоферовой) зоне распределения источников, то формулы (9.5) и (9.8) упрощаются, так как можно воспользоваться приближенным выражением функции Грина. Пусть начало координат помещено в центр области, занятой источниками. Тогда при  $r \gg ka^2$  расстояние  $R = |\mathbf{r} - \mathbf{r}'|$  в формуле (9.9) можно приближенно заменить на  $r - (\mathbf{n}\mathbf{r}')$  в показателе экспоненты и на  $r$  в знаменателе, что и приводит к фраунгоферову приближению. В этом приближении (9.8) принимает вид

$$u(\omega, \mathbf{r}) \approx -\frac{e^{ikr}}{4\pi r} \int q(\omega, \mathbf{r}') e^{-ik\mathbf{n}\mathbf{r}'} d^3r', \quad r \gg ka^2, \quad (9.10)$$

где  $\mathbf{n} = \mathbf{r}/r$  — единичный вектор в направлении на точку наблюдения.



Приведем теперь динамические соотношения для того случая, когда заданы *виртуальные* источники. Будем исходить из формулы Грина, которая связывает спектральную амплитуду поля  $u(\omega, \mathbf{r})$  внутри области, ограниченной поверхностью  $S_0$ , со спектральной амплитудой граничного поля  $v(\omega, \mathbf{r}')$ :

$$u(\omega, \mathbf{r}) = -\frac{1}{4\pi} \oint \left[ v(\omega, \mathbf{r}') \frac{\partial}{\partial N} \left( \frac{e^{ikR}}{R} \right) - \frac{e^{ikR}}{R} \frac{\partial v(\omega, \mathbf{r}')}{\partial N} \right] dS_0. \quad (9.11)$$

Здесь  $R = |\mathbf{r} - \mathbf{r}'|$  — расстояние от точки наблюдения до точки  $\mathbf{r}'$ , лежащей на поверхности  $S_0$ , по которой ведется интегрирование, а  $\partial/\partial N$  означает дифференцирование в направлении внешней нормали  $\mathbf{N}$  к  $S_0$ . Граничные значения  $v(\omega, \mathbf{r}')$  и  $\partial v(\omega, \mathbf{r}')/\partial N$ , как известно, должны быть заданы математически непротиворечивым образом. Поэтому двучленная формула Грина (9.11) может быть использована лишь в тех случаях, когда из каких-либо дополнительных соображений вытекает связь (точная или приближенная) между граничным полем  $v$  и его нормальной производной  $\partial v/\partial N$ .

Особым является случай, когда поверхностью  $S_0$  служит плоскость, замкнутая полусферой бесконечно большого радиуса. В этом случае, которым мы и ограничимся, двучленная формула (9.11) сводится к одночленной.

Если на плоскости  $z=0$  задано само поле  $v(\omega, \mathbf{p}) \equiv u(\omega, \mathbf{r})|_{z=0}$ , где  $\mathbf{p} = (x, y)$  — двумерный вектор, то интеграл по бесконечно удаленной сфере обращается в нуль, и тогда в полупространстве  $z > 0$

$$u(\omega, \mathbf{r}) = u(\omega, \mathbf{p}, z) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} v(\omega, \mathbf{p}') \frac{\partial}{\partial z} \left( \frac{e^{ikR}}{R} \right) d^2\mathbf{p}', \quad (9.12)$$

$$R = \sqrt{(\mathbf{p} - \mathbf{p}')^2 + z^2}.$$

Если же на плоскости  $z=0$  задана нормальная производная  $\partial v(\omega, \mathbf{p})/\partial z = \partial u(\omega, \mathbf{r})/\partial z|_{z=0}$ , то

$$u(\omega, \mathbf{p}, z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{ikR}}{R} \frac{\partial v(\omega, \mathbf{p}')}{\partial z} d^2\mathbf{p}'. \quad (9.13)$$

Таким образом, в случае плоской границы  $z=0$  достаточно задания на  $S_0$  либо самого поля  $u$ , либо его нормальной производной  $\partial u/\partial z$ .

Другой метод расчета полей в полупространстве  $z > 0$  восходит к Релею и основан на разложении полей по плоским волнам. Представим граничное поле  $v(\omega, \mathbf{p})$  двумерным

интегралом Фурье:

$$v(\omega, \rho) = \int_{-\infty}^{\infty} v(\omega, \kappa) \exp(i\kappa\rho) d^2\kappa, \quad (9.14)$$

где  $\kappa = (\kappa_x, \kappa_y)$  — двумерный волновой вектор, а  $v(\omega, \kappa)$  —  $\omega\kappa$ -амплитуда граничного поля, связанная с  $v(\omega, \rho)$  обратным преобразованием Фурье:

$$v(\omega, \kappa) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} v(\omega, \rho) \exp(-i\kappa\rho) d^2\rho.$$

В полупространстве  $z > 0$  каждая пространственная гармоника граничного поля  $v(\omega, \kappa) \exp(i\kappa\rho)$  порождает плоскую собственную волну  $v(\omega, \kappa) \exp(i\kappa\rho + ipz)$  (напомним, что множитель  $e^{-i\omega t}$  опущен), которая удовлетворяет волновому уравнению при условии, что

$$\kappa^2 + p^2 = \omega^2/c^2 \equiv k^2.$$

Таким образом, в случае волнового уравнения для однородной и изотропной среды дисперсионная поверхность (см. § 6) представляет собой в четырехмерном пространстве  $(\omega, \mathbf{k})$  трехмерный конус с осью по  $\omega$  и вершиной в начале координат. В соответствии с этим дисперсионным уравнением в свободной от источников области  $z > 0$  возможны при заданном  $\kappa$  собственные волны двух типов — в зависимости от того, вещественна или мнима  $z$ -компонента  $p$  волнового вектора  $\mathbf{k}$ :

$$p = \sqrt{k^2 - \kappa^2} = \begin{cases} \sqrt{k^2 - \kappa^2} & \kappa \leq k = \omega/c, \\ i\sqrt{|\kappa^2 - k^2|} & \kappa > k. \end{cases} \quad (9.15)$$

При  $\kappa < k$  компонента  $p$  вещественна и мы имеем бегущие волны. При  $\kappa > k$ , когда период осцилляций на границе  $2\pi/\kappa$  меньше длины волны  $\lambda = 2\pi/k$ , соответствующей частоте  $\omega$ , компонента  $p$  мнима и получаются неоднородные волны. Они экспоненциально ослабевают с удалением от границы  $z = 0$ . Практически уже при  $z \geq \lambda$  остаются только бегущие волны.

Результирующее волновое поле  $u(\omega, \mathbf{r}) = u(\omega, \rho, z)$ , удовлетворяющее граничному условию  $u(\omega, \mathbf{r})|_{z=0} = v(\omega, \rho)$ , выражается суперпозицией плоских волн обоих типов:

$$u(\omega, \rho, z) = \int_{-\infty}^{\infty} v(\omega, \kappa) \exp(i\kappa\rho + ipz) d^2\kappa. \quad (9.16)$$

Аналогичное представление поля нетрудно записать и в том случае, когда на плоскости  $z = 0$  задана производная по нормали  $du/dz$ .

Формула Грина (9.12) и разложение по плоским волнам (9.16), разумеется, эквивалентны друг другу. В дальнейшем мы будем пользоваться той из них, которая быстрее ведет к окончательному результату. Любая из них позволяет выразить моменты  $u(\omega, \mathbf{r})$  через моменты  $v(\omega, \rho)$ .

При вычислении моментов поля при помощи приведенных выше формул часто приходится рассматривать те или иные частные случаи. Перечислим наиболее существенные из них. При  $kz \gg 1$ , т. е. в волновой зоне, для ядра в формуле (9.12) имеем

$$-\frac{1}{2\pi} \frac{\partial}{\partial z} \frac{e^{ikR}}{R} \approx \frac{kz}{2\pi i R^2} e^{ikR} \quad (kz \gg 1).$$

При выполнении неравенств  $|\rho - \rho'| \ll z$  и  $k|\rho - \rho'|^4 \ll z^3$  справедливо так называемое *френелевское приближение*, в котором  $R = \sqrt{z^2 + (\rho - \rho')^2}$  заменяется на  $z + (\rho - \rho')^2/2z$  в показателе экспоненты и на  $z$  в знаменателе предыдущей формулы; при этом формула Грина (9.12) принимает вид

$$u(\omega, \rho, z) = \frac{ke^{ikz}}{2\pi iz} \int_{-\infty}^{\infty} v(\omega, \rho') \exp\left(ik \frac{(\rho - \rho')^2}{2z}\right) d^2\rho'. \quad (9.17)$$

Наконец, в зоне *Фраунгофера*,  $z \gg ka^2$ , где  $a$  — радиус области, в которой граничное поле  $v(\omega, \rho)$  отлично от нуля, имеем

$$u(\omega, \rho, z) = \frac{k \exp[ik(z + \rho^2/2z)]}{2\pi iz} \int_{-\infty}^{\infty} v(\omega, \rho') \exp\left[-ik \frac{(\rho\rho')}{z}\right] d^2\rho'. \quad (9.18)$$

Обратимся теперь к некоторым свойствам моментов и спектров волновых полей. Статистические моменты волновых случайных полей часто называют *функциями когерентности*, так как соответствующие коэффициенты корреляции служат количественной мерой когерентности этих полей (ч. I, §§ 47 и 48). Термины «время когерентности», «длина когерентности», «степень когерентности» являются в отношении волновых полей синонимами «времени корреляции», «радиуса корреляции», «коэффициента корреляции». В связи с этим мы предпочитаем и в этой части книги говорить о *теории случайных волн*, а не о «теории когерентности», рассматривая последнюю просто как одно из приложений общей теории случайных полей. Однако мы не будем избегать и терминов «функция когерентности» или «степень когерентности», которые уже прочно вошли в физический обиход.

Смешанный момент волнового поля  $u(t, \mathbf{r})$  порядка  $m+n$  (функция когерентности порядка  $m+n$ ) определяется соотношением вида (2.30) и вместо  $B_{m,n}$  обычно обозначается  $\Gamma_{m,n}$ . Корреляционная теория случайных волн ограничивается рассмотрением

моментов лишь первого и второго порядков: *среднего поля*  $\langle u \rangle$  (функции когерентности первого порядка  $\Gamma_1 \equiv \langle u \rangle$ ) и двух *смешанных моментов второго порядка*, называемых первой и второй функциями когерентности второго порядка:

$$\begin{aligned}\Gamma_{1,1}(1,2) &= \langle u(1) u^*(2) \rangle \equiv \Gamma(1,2), \\ \Gamma_{2,0}(1,2) &= \langle u(1) u(2) \rangle \equiv \tilde{\Gamma}(1,2).\end{aligned}\quad (9.19)$$

Последние связаны с первой и второй функциями корреляции поля формулами

$$\begin{aligned}\psi(1,2) &= \langle \tilde{u}(1) \tilde{u}^*(2) \rangle = \Gamma(1,2) - \langle u(1) \rangle \langle u^*(2) \rangle, \\ \tilde{\psi}(1,2) &= \langle \tilde{u}(1) \tilde{u}(2) \rangle = \tilde{\Gamma}(1,2) - \langle u(1) \rangle \langle u(2) \rangle\end{aligned}\quad (9.20)$$

(как обычно, здесь  $\tilde{u} = u - \langle u \rangle$ ).

Для зависимости волновых полей от времени мы примем комплексное представление в виде аналитического сигнала<sup>1)</sup>

$$u(t, \mathbf{r}) = \int_0^\infty u(\omega, \mathbf{r}) e^{-i\omega t} d\omega, \quad (9.21)$$

где спектральная амплитуда  $u(\omega, \mathbf{r})$  тождественно равна нулю для отрицательных частот, а для  $\omega \geq 0$  равна удвоенной спектральной плотности исходного вещественного поля  $u_r(t, \mathbf{r}) = \text{Re } u(t, \mathbf{r})$ .

Если поле  $u(t, \mathbf{r})$  является как функция  $t$  аналитическим сигналом и *стационарно*, то среднее значение поля  $\langle u \rangle$  и его второй смешанный момент  $\tilde{\Gamma}$  равны нулю (ч. I, § 38).

Без предположения об аналитическом сигнале среднее значение стационарного поля может быть отличным от нуля и представляет собой некоторую функцию ~~только~~ от  $\mathbf{r}$ , удовлетворяющую уравнению Лапласа  $\Delta u(\mathbf{r}) = 0$ . Так как такие статистические детерминированные поля нас не интересуют, можно и в этом случае рассматривать только флуктуации  $\tilde{u} = u - \langle u \rangle$ , т. е. считать, что  $\langle \tilde{u} \rangle = 0$ . Что касается нестационарных и, в частности, монохроматических полей ( $\sim e^{-i\omega t}$ ), то для них среднее поле и  $\tilde{\Gamma}$ , вообще говоря, отличны от нуля.

Энергетические величины (интенсивность, плотность энергии, плотность ее потока) квадратичны по полю, в силу чего их средние значения можно выразить через статистические моменты первого и второго порядков (ч. I, § 47 и задача I).

<sup>1)</sup> Еще раз напомним, что при разложении в интеграл Фурье по  $t$  мы берем здесь и далее функцию  $e^{-i\omega t}$ , а не  $e^{i\omega t}$ , как в ч. I. Поэтому поле  $u$  как функция  $t$  аналитично теперь не в верхней, а в нижней полуплоскости комплексного времени  $t$ .

Из высших моментов наибольший интерес представляет момент четвертого порядка

$$\Gamma_{2,2}(1, 2, 3, 4) = \langle u(1) u(2) u^*(3) u^*(4) \rangle, \quad (9.22)$$

через который выражается, в частности, корреляционная функция *флуктуаций интенсивности* комплексного поля  $I = |u|^2$ :

$$\psi_I(1, 2) = \langle \tilde{I}(1) \tilde{I}(2) \rangle = \Gamma_{2,2}(1, 2; 2, 1) - \bar{I}(1) \bar{I}(2), \quad (9.23)$$

причем  $\bar{I}(1) = \langle |u(1)|^2 \rangle = \Gamma(1, 1)$ . Если поле  $u$  нормально, то все высшие моменты выражаются через первый ( $u$ ) и вторые ( $\Gamma$  и  $\tilde{\Gamma}$ ) моменты. Функция корреляции интенсивности для этого случая приведена (в несколько иных обозначениях) в задаче 12 к гл. I.

Выражения для моментов поля мы получали выше, используя динамические решения для  $u$ . Как уже было отмечено, существует и другой способ нахождения моментов — из уравнений, которым подчиняются сами статистические моменты. В рассматриваемой задаче о возбуждении полей реальными источниками, которая в общем случае описывается уравнением (8.1) с детерминированным оператором  $\hat{L}$  и с детерминированными граничными условиями, уравнения для вторых моментов легко получить простым перемножением левых и правых частей (8.1), взятых в разных пространственно-временных точках. Так, второй момент  $\Gamma(1, 2) = \langle u(1) u^*(2) \rangle$  удовлетворяет уравнению

$$\hat{L}_1 \hat{L}_2^* \Gamma(1, 2) = \langle q(1) q^*(2) \rangle, \quad (9.24)$$

которое становится однородным в области, свободной от источников.

В рамках схемы 1), когда динамическое решение задачи для поля  $u$  известно, находить моменты из уравнений типа (9.24) обычно гораздо менее удобно, чем по формулам типа (9.1) или (9.3). Волновые уравнения для моментов представляют здесь интерес, пожалуй, лишь в том отношении, что из них очевидны *волновые свойства самих моментов* и, соответственно, можно говорить о распространении и дифракции этих моментов почти в том же смысле, что и для поля  $u$ <sup>1)</sup>. Одно из важных волновых свойств статистических моментов поля  $u$  состоит в том, что их значения на некоторой поверхности  $S_0$  определяют поведение моментов во всем объеме, ограниченном поверхностью  $S_0$ , подобно тому как это имеет место для самого поля.

<sup>1)</sup> Однако в тех случаях, когда мы не располагаем решением динамической задачи, уравнения для моментов, получаемые из уравнений для  $u$ , приобретают самостоятельную ценность (гл. VI—VIII).

Как уже было отмечено в § 6, в рассматриваемом случае волнового уравнения для однородной и стационарной среды дисперсионная гиперповерхность представляет собой конус  $\kappa^2 = \omega^2/c^2 \equiv k^2$  или  $|\kappa| = k$ . Поэтому спектральная плотность  $G(\omega, \kappa)$  в разложении Фурье корреляционной функции однородного и стационарного поля (6.3) содержит множитель  $\delta(|\kappa| - k)$ :

$$G(\omega, \kappa) = f(\omega, n) \delta(|\kappa| - k), \quad (9.25)$$

где  $f(\omega, n)$  — функция частоты  $\omega$  и направления  $n = \kappa/\kappa$ .

В коэффициенте при дельта-функции можно выделить множитель  $k^{-2}$ :

$$G(\omega, \kappa) = \mathcal{I}(\omega, n) k^{-2} \delta(\kappa - k). \quad (9.26)$$

Величина  $\mathcal{I}(\omega, n)$  носит название *лучевой интенсивности* или *яркости* и играет большую роль в теории переноса излучения (гл. VIII). Связь этой величины с функцией корреляции можно установить, подставив (9.26) в (6.3) и положив  $d^3\kappa = \kappa^2 d\kappa d\Omega(n)$  ( $d\Omega(n)$  — элемент телесного угла). Выполнив интегрирование по  $\kappa$ , получаем искомое соотношение ( $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ ,  $t = t_1 - t_2$ ):

$$\begin{aligned} \Psi(t, \mathbf{r}) &= \iint \mathcal{I}(\omega, n) k^{-2} \delta(\kappa - k) \exp[i(\kappa \mathbf{r} - \omega t)] \kappa^2 d\kappa d\Omega(n) d\omega = \\ &= \int_0^\infty d\omega \oint \mathcal{I}(\omega, n) \exp[i(kn\mathbf{r} - \omega t)] d\Omega(n). \end{aligned} \quad (9.27)$$

При  $\mathbf{r} = 0$  и  $t = 0$  эта формула дает дисперсию поля

$$\sigma_u^2 = \langle |\tilde{u}|^2 \rangle = \Psi(0, 0) = \int_0^\infty d\omega \oint \mathcal{I}(\omega, n) d\Omega(n). \quad (9.28)$$

Из последнего выражения видно, что лучевая интенсивность  $\mathcal{I}(\omega, n)$  описывает распределение энергии по (положительным) частотам и по углам, т. е. представляет собой *частотно-угловой спектр* поля. В частности, изотропному полю отвечает лучевая интенсивность, не зависящая от углов (общее выражение для функции корреляции в этом случае дано в задаче 2). Заметим, что угловым спектром часто называют также двумерное разложение функции корреляции в интеграл Фурье:

$$F(\omega; \kappa_\perp, z) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \Psi(\omega; \rho, z) \exp(-i\kappa_\perp \rho) d^2\rho, \quad \kappa_\perp = (\kappa_x, \kappa_y). \quad (9.29)$$

Как показано в задаче 4, при малой ширине углового спектра величина  $F(\omega; \kappa_\perp, z)$  пропорциональна лучевой интенсивности (при  $z = 0$  эти величины отличаются лишь множителем  $k^{-2}$ ).

При помощи (9.27) можно оценить минимальный радиус корреляции  $l_n$  свободного волнового поля. Пусть  $\omega_{\max}$  — максимальная частота волн, еще заметно представленных в спектре  $\mathcal{J}(\omega, n)$ . В силу (9.27) радиус корреляции поля  $l_n$  можно оценить из условия  $k_{\max}(n, r) \geq 1$ , что приводит к оценке

$$l_n \geq \frac{1}{k_{\max}} \sim \lambda_{\min}.$$

Иными словами, радиус корреляции случайного волнового поля не может быть меньше минимальной длины волны, имеющейся в спектре колебаний. Строго говоря, эта оценка относится к случаю изотропного поля, когда лучевая интенсивность не зависит от  $n$ . Оценки продольного и поперечного радиусов корреляции поля для случая, когда его лучевая интенсивность сосредоточена в узком конусе, даны в задаче 3.

Описанные выше свойства статистических моментов и спектральных плотностей характерны не только для скалярных, но и для векторных случайных полей. Мы не будем приводить здесь векторные аналоги рассмотренных выражений для полей и их моментов, поскольку в принципе они не выходят за рамки общих соотношений (9.1) и (9.3). Отметим только, что статистические характеристики поляризации плоских квазимонохроматических волн мы уже рассмотрели в ч. I, § 49. Введенная там матрица поляризации характеризует статистическую связь разных компонент поля в одной и той же точке в один и тот же момент времени (см. также [13, 14, 21] и [22]). Общий вид корреляционной матрицы стационарного, однородного и изотропного поля приведен в задаче 6, а в задаче 5 указана связь энергетических характеристик поля с корреляционной матрицей поля и с матрицей лучевых интенсивностей.

## § 10. Дифракция плоской волны на безграничном хаотическом экране

Первичное поле на границе рассматриваемой области может флуктуировать как из-за флуктуаций в его источниках (находящихся вне данной области), так и в результате случайных возмущений, внесенных в первоначально детерминированную первичную волну при ее распространении. Примером может служить прохождение этой волны через случайно-неоднородную среду или слой такой среды. Если толщина этого слоя достаточно мала, то его можно рассматривать как бесконечно тонкий экран. Применимость общих методов расчета поля в области, на границе которой  $S_0$  задано случайное поле (§ 9), конечно, не связана с тем, по каким причинам флуктурует поле на границе. Существенно лишь то, что статистика этих флуктуаций известна.

Начнем с простейшей задачи о прохождении плоской монохроматической волны через плоский безграничный хаотический экран, следуя в основном работе [23]<sup>1)</sup>.

1. Основные соотношения. Пусть монохроматическая волна  $u_n(\rho, z)e^{-i\omega t}$  падает на безграничный экран, расположенный в плоскости  $z=0$ . Экран пространственно модулирует падающую волну в соответствии со своей функцией пропускания  $f(\rho)$  — комплексной функцией, модуль которой описывает амплитудную модуляцию, а аргумент — фазовую. Граничное поле  $v(\rho)$ , т. е. поле непосредственно за экраном  $z=0$ , есть

$$v(\rho) = f(\rho) u_n(\rho, 0). \quad (10.1)$$

Для краткости мы опускаем множитель  $e^{-i\omega t}$ , а также аргумент  $\omega$  в спектральных амплитудах.

Если падающая волна плоская и распространяется по нормали к экрану, то  $u_n(\rho, z) = e^{ikz}$ , так что

$$v(\rho) = f(\rho) \quad (10.2)$$

— поле на граничной плоскости  $z=0$  просто равно функции пропускания, и, соответственно, статистика  $v(\rho)$  в случае хаотического экрана та же, что у  $f(\rho)$ . Формулы (9.12) и (9.16) позволяют связать статистические моменты поля  $u(\mathbf{r}) = u(\rho, z)$  за экраном (т. е. в области  $z > 0$ ) с моментами граничного поля  $v(\rho)$ . Найдем первый и второй моменты поля  $u(\mathbf{r})$  в простейшем случае статистически однородного хаотического экрана, для которого

$$\langle f(\rho) \rangle = f_0 = \text{const}, \quad \sigma_f^2 = \langle |f|^2 \rangle = \langle |f - f_0|^2 \rangle = \text{const},$$

а корреляционная функция зависит только от разностей координат:

$$\psi_f(\rho', \rho'') = \langle f(\rho') f^*(\rho'') \rangle = \psi_f(\rho' - \rho'').$$

В силу (10.2) имеем

$$\langle v \rangle = v_0 = f_0, \quad \sigma_v^2 = \sigma_f^2, \quad \psi_v(\rho' - \rho'') = \psi_f(\rho' - \rho'').$$

Для среднего поля за экраном по формуле (9.12) находим

$$\langle u(\rho, z) \rangle = -\frac{v_0}{2\pi} \int_{-\infty}^{\infty} \frac{\partial}{\partial z} \frac{e^{ikR}}{R} d^2\rho'.$$

<sup>1)</sup> В этой работе суммированы результаты и предшествующих исследований [24—26], относящихся преимущественно к прохождению радиоволн через случайно-неоднородную ионосферу. Более полный список литературы по последнему вопросу содержится в обзоре [27].



Интеграл легко вычисляется и равен  $-2\pi \exp(ikz)$ . Таким образом, среднее поле за экраном — это плоская волна

$$\bar{u} = v_0 e^{ikz}, \quad (10.3)$$

амплитуда которой  $v_0$  равна средней прозрачности  $f_0$ .

Для нахождения функции корреляции воспользуемся выражением (9.16), из которого следует, что

$$\begin{aligned} \psi_u(\rho_1, z_1; \rho_2, z_2) &= \langle \tilde{u}(\mathbf{r}_1) \tilde{u}^*(\mathbf{r}_2) \rangle = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d^2\kappa' d^2\kappa'' \langle \tilde{v}(\kappa') \tilde{v}^*(\kappa'') \rangle \exp(i\kappa' \rho_1 - i\kappa'' \rho_2 + ip' z_1 - ip'' z_2), \end{aligned} \quad (10.4)$$

где  $p'$  и  $p''$  даются выражением (9.15) соответственно при  $\kappa = \kappa'$  и  $\kappa = \kappa''$ . Но для однородного граничного поля

$$\langle \tilde{v}(\kappa') \tilde{v}^*(\kappa'') \rangle = F_v(\kappa') \delta(\kappa' - \kappa''), \quad (10.5)$$

где  $F_v(\kappa)$  — двумерная спектральная плотность, через которую выражается корреляционная функция граничного поля:

$$\psi_v(\rho) = \int_{-\infty}^{\infty} F_v(\kappa) \exp(i\kappa \rho) d^2\kappa. \quad (10.6)$$

Подставляя (10.5) в (10.4), находим

$$\psi_u = \psi_u(\rho_1 - \rho_2, z_1, z_2) = \int_{-\infty}^{\infty} F_v(\kappa) \exp[i\kappa(\rho_1 - \rho_2) + i(pz_1 - p^* z_2)] d^2\kappa. \quad (10.7)$$

Таким образом, поле  $u(\rho, z)$  статистически однородно в плоскостях  $z = \text{const}$ , что является следствием однородности  $v(\rho)$  в плоскости  $z = 0$ . Из сравнения (10.7) с двумерным спектральным представлением

$$\psi_u(\rho_1 - \rho_2, z_1, z_2) = \int_{-\infty}^{\infty} F_u(\kappa, z_1, z_2) \exp[i\kappa(\rho_1 - \rho_2)] d^2\kappa \quad (10.8)$$

следует, что пространственная спектральная плотность есть

$$\begin{aligned} F_u(\kappa; z_1, z_2) &= F_v(\kappa) \exp[i(pz_1 - p^* z_2)] = \\ &= F_v(\kappa) \times \begin{cases} \exp[i\sqrt{k^2 - \kappa^2}(z_1 - z_2)], & \kappa^2 \leq k^2, \\ \exp[-\sqrt{|\kappa^2 - k^2|}(z_1 + z_2)], & \kappa^2 > k^2. \end{cases} \end{aligned}$$

Спектральные составляющие с  $\kappa > k$  (неоднородности граничного поля  $v(\rho)$  меньше длины волны) порождают в полупространстве  $z > 0$  неоднородные волны, ослабевающие при удалении от экрана по экспоненциальному закону. Для точек наблюдения  $z_1$

и  $z_2$ , удаленных от экрана уже на несколько длин волн ( $z_1, z_2 \gg \lambda$ ), приближенно

$$F_u(\kappa; z_1, z_2) \approx \begin{cases} F_v(\kappa) \exp[i\sqrt{k^2 - \kappa^2}(z_1 - z_2)], & \kappa^2 \leq k^2, \\ 0, & \kappa^2 > k^2. \end{cases} \quad (10.9)$$

Подставляя (10.9) в (10.8), получаем для корреляционной функции выражение

$$\psi_u(\rho, \xi) = \int_{\kappa \leq k} F_v(\kappa) \exp[i\kappa\rho + i\sqrt{k^2 - \kappa^2}\xi] d^2\kappa, \quad (10.10)$$

где  $\rho = \rho_1 - \rho_2$ ,  $\xi = z_1 - z_2$ . Из (10.10) видно, что при  $z_1, z_2 \gg \lambda$  случайное поле  $u(\rho, z)$  становится статистически однородным не только в поперечных плоскостях  $z = \text{const}$ , но и в продольном направлении  $z$ .

Используя формулу (10.10), рассмотрим частные случаи мелкомасштабных и крупномасштабных флуктуаций граничного поля  $v(\rho)$ .

В случае *мелкомасштабных* флуктуаций, когда радиус корреляции граничного поля  $l_v$  мал по сравнению с длиной волны  $\lambda$ , т. е. при  $kl_v \ll 1$ , можно принять, что спектральная плотность  $F_v(\kappa)$  практически постоянна в круге  $\kappa \leq k$  и приближенно равна  $F_v(0) \equiv F_0$ . Тогда из (10.10) получаем

$$\psi_u(\rho, \xi) = F_0 \int_{\kappa \leq k} \exp(i\kappa\rho + i\rho\xi) d^2\kappa, \quad kl_v \ll 1. \quad (10.11)$$

*Поперечная* функция корреляции, т. е. функция корреляции в плоскости  $z = \text{const}$ , перпендикулярной к направлению распространения, получается из (10.11) при  $\xi = z_1 - z_2 = 0$ :

$$\psi_{\perp}(\rho) \equiv \psi_u(\rho, 0) = F_0 \int_{\kappa \leq k} \exp(i\kappa\rho) d^2\kappa = 2\pi F_0 \frac{k J_1(k\rho)}{\rho},$$

где  $J_1$  — функция Бесселя первого порядка. При увеличении  $\rho$  отношение  $J_1(k\rho)/\rho$  проходит в первый раз через нуль при  $k\rho = 3,83$ , т. е. при  $\rho \approx 0,61\lambda$ . Таким образом, уже на расстояниях в несколько  $\lambda$  от экрана поперечный радиус корреляции поля  $l_{\perp}$  — величина порядка длины волны, т. е. значительно больше радиуса корреляции  $l_v$  в плоскости  $z = 0$ :  $l_{\perp} \sim \lambda \gg l_v$ .

Из выражения для *продольной* функции корреляции

$$\psi_{\parallel}(\xi) \equiv \psi_u(0, \xi) = F_0 \int_{\kappa \leq k} e^{i\rho\xi} d^2\kappa = 2\pi F_0 \frac{d}{d\xi} \left( \frac{1 - e^{-ik\xi}}{\xi} \right)$$

следует, что продольный радиус корреляции  $l_{\parallel}$  тоже порядка длины волны:  $l_{\parallel} \sim \lambda \gg l_v$ . Таким образом, в случае мелкомасштабных флуктуаций на экране как поперечный, так и продоль-

ный радиусы корреляции поля при удалении от экрана увеличиваются и при  $z \gg \lambda$  достигают значений порядка длины волны.

Вместе с тем происходит уменьшение дисперсии флуктуаций, т. е. их сглаживание. Действительно, при помощи (10.11) находим

$$\sigma_u^2 = \langle |\tilde{u}|^2 \rangle = \psi_u(0, 0) = F_0 \int_{\kappa \leq k} d^2\kappa = \pi k^2 F_0.$$

Но

$$F_0 \equiv F_v(0) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \psi_v(\rho') d^2\rho' \sim \frac{1}{4\pi^2} \sigma_v^{2/2},$$

так что

$$\sigma_u^2 = \pi k^2 F_0 \sim (kl_v)^2 \sigma_v^2 \ll \sigma_v^2.$$

В пределе при  $kl_v \rightarrow 0$  (очень мелкие неоднородности экрана) флуктуационное поле  $\tilde{u}$  вообще исчезает, так как столь мелкие неоднородности порождают за экраном только неоднородные (экспоненциально спадающие) волны.

Обратимся теперь к *крупномасштабным* неоднородностям граничного поля ( $kl_v \gg 1$ ). Такие флуктуации порождают бегущие волны, в силу чего они особенно важны в приложениях. Мы уделим им поэтому основное внимание.

В случае крупномасштабных флуктуаций граничного поля двумерный спектр  $F_v(\kappa)$  сосредоточен в узком интервале значений  $\kappa \leq \kappa_v \sim 1/l_v \ll k$ . Это позволяет разложить продольное волновое число  $p = \sqrt{k^2 - \kappa^2}$  в формуле (10.10) в ряд Тейлора по степеням  $\kappa^2$  и ограничиться двумя первыми членами:

$$p = \sqrt{k^2 - \kappa^2} \approx k - \kappa^2/2k.$$

Кроме того, не совершая заметной ошибки, можно раздвинуть пределы интегрирования в (10.10) до  $\pm\infty$ , и тогда

$$\psi_u(\rho, \xi) \approx e^{ik\xi} \int_{-\infty}^{\infty} F_v(\kappa) \exp(i\kappa\rho - i\kappa^2\xi/2k) d^2\kappa \quad (kl_v \gg 1). \quad (10.12)$$

Нетрудно убедиться, что сделанные допущения эквивалентны использованию для поля  $u(\rho, z)$  френелевского приближения (9.17).

Положив в (10.12)  $\xi = 0$ , получаем *поперечную* функцию корреляции:

$$\psi_{\perp}(\rho) \equiv \psi_u(\rho, 0) = \int_{-\infty}^{\infty} F_v(\kappa) \exp(i\kappa\rho) d^2\kappa = \psi_v(\rho) \quad (10.13)$$

(см. (10.6)). Таким образом, поперечная функция корреляции волнового поля  $u$  равна функции корреляции граничного поля

и не меняется при удалении от экрана. Тем самым и поперечный радиус корреляции такой же, как в плоскости экрана:

$$l_{\perp} = l_v.$$

Сохранение поперечной функции корреляции, отмеченное впервые в работе [24], означает сохранение и дисперсии ( $\sigma_u^2 = \sigma_v^2$ ), и поперечной функции когерентности. В самом деле, учитывая закон изменения среднего поля (10.3), имеем для волны, не ограниченной в поперечной плоскости,

$$\begin{aligned} \Gamma_{\perp}(\rho, z) &= \langle u(\rho_1, z) u^*(\rho_2, z) \rangle = \psi_{\perp}(\rho) + \bar{u}(\rho_1, z) \bar{u}^*(\rho_2, z) = \\ &= \psi_v(\rho) + |v_0|^2 = \Gamma_v(\rho). \end{aligned} \quad (10.14)$$

При  $\rho = \rho_1 - \rho_2 = 0$  отсюда вытекает постоянство средней интенсивности при удалении от экрана:

$$\bar{I}_u(z) = \langle |u|^2 \rangle = \Gamma_v(0) = \langle |v|^2 \rangle = \text{const.}$$

Следует подчеркнуть, что сохранение  $\psi_{\perp}$ ,  $\Gamma_{\perp}$  и  $\bar{I}$  при удалении от плоскости  $z=0$  не имеет места для статистически неоднородного экрана и для неплоской падающей волны, а более высокие моменты поля не сохраняются даже для плоской волны и статистически однородного экрана.

Продольная функция корреляции  $\psi_{\parallel}(\xi)$  получается из (10.12) при  $\rho=0$ :

$$\psi_{\parallel}(\xi) = e^{ik\xi} \int_{-\infty}^{\infty} F_v(k) e^{-ik^2\xi/2k} d^2k.$$

Заметное уменьшение модуля функции  $\psi_{\parallel}(\xi)$  по сравнению с максимальным значением  $\psi_{\parallel}(0) = \sigma_u^2$  наступает при  $\xi \sim 2k/\kappa_v^2$ , когда подынтегральная экспонента начинает осциллировать в пределах интервала  $\kappa \leq \kappa_v \sim 1/l_v$ , в котором сосредоточен спектр  $F_v(k)$ . Отсюда можно оценить продольный радиус корреляции  $l_{\parallel}$ :

$$l_{\parallel} \sim \frac{2k}{\kappa_v^2} \sim kl_v^2,$$

т. е.  $l_{\parallel}$  в  $kl_v \gg 1$  раз больше поперечного масштаба  $l_{\perp} \sim l_v$ . Можно сказать, что продольная корреляция флуктуаций поля исчезает тогда, когда радиус первой зоны Френеля  $\sqrt{\lambda|z_1 - z_2|} = \sqrt{\lambda|\xi|}$  для отрезка  $|\xi| = |z_1 - z_2|$  становится больше поперечного радиуса корреляции:  $\sqrt{\lambda l_{\parallel}} \sim l_v$ .

Проиллюстрируем сказанное о соотношении между продольным и поперечным масштабами корреляции примером, в котором функция корреляции поля вычисляется точно. Пусть у поля на экране корреляционная функция гауссова:  $\psi_v(\rho) = \sigma_v^2 \exp[-\rho^2/2l_v^2]$ .

Для нее

$$F_{\gamma}(\kappa) = \sigma_v^2 \exp[-\kappa^2 l_v^2/2],$$

и простой расчет по формуле (10.12) дает

$$\psi_{\kappa}(\rho, \zeta) = \sigma_v^2 \gamma e^{i\kappa \zeta} \exp[-\rho^2 \gamma/2 l_v^2] = \gamma \psi_v(\rho) \exp\{(1-\gamma)\rho^2/2 l_v^2\},$$

где  $\gamma = [1 + i\zeta/k l_v^2]^{-1}$ . Характерный масштаб изменения  $|\psi_{\kappa}|$  имеет порядок  $l_{\perp} \sim l_v$  в поперечном направлении и  $l_{\parallel} \sim k l_v^2 \gg l_v$  в продольном. Поверхности равных значений  $|\psi_{\kappa}|$  имеют при этом форму овалов, сильно вытянутых вдоль оси  $z$ .

2. Прохождение плоской волны через фазовый хаотический экран. Так называют прозрачный экран с функцией пропускания  $f(\rho) = \exp[iS(\rho)]$ , где  $S$  — вещественная случайная функция, т. е. экран модулирует только фазу, но оставляет неизменной амплитуду (и, следовательно, интенсивность) волны. В полупространстве  $z > 0$  фазовый экран вызывает в прошедшей волне как фазовую, так и амплитудную модуляцию. Последнюю можно наблюдать, например, на листе бумаги при прохождении света через оптически неоднородное или неровное (скажем, обычное оконное) стекло, если отодвинуть бумагу на некоторое расстояние от стекла.

Фазовый экран часто используется в качестве модели для описания ряда явлений как в оптике, так и в радиофизике. Например, линза с оптическими неоднородностями модулирует главным образом фазу проходящей световой волны. Такое же действие оказывает ионосфера Земли на проходящие через нее радиоволны УКВ-диапазона. Модель фазового экрана применяют также (хотя и с меньшими основаниями) при анализе мерцаний радиоволн, посылаемых внеземными радиоисточниками и проходящих через статистически неоднородную межпланетную или межзвездную среду.

Если на неограниченный плоский фазовый экран падает плоская волна  $u_n = e^{ikz}$ , то граничное значение поля в плоскости экрана  $z=0$  равно

$$v(\rho) = \exp[iS(\rho)] u_n|_{z=0} = \exp[iS(\rho)].$$

Выясним, как связаны статистические моменты поля прошедшей волны с функцией корреляции фазы

$$\psi_S(\rho) = \langle \tilde{S}(\rho_1) \tilde{S}(\rho_2) \rangle = \sigma_S^2 K_S(\rho),$$

где  $\sigma_S^2$  — дисперсия, а  $K_S(\rho)$  — коэффициент корреляции фазы.

Предположим, что флуктуации  $S$  статистически однородны в плоскости  $z=0$ , имеют нулевое среднее значение ( $\bar{S}=0$ ) и подчиняются нормальному закону распределения вероятностей.

Учитывая, что для нормально распределенной величины  $\alpha$  с  $\bar{\alpha}=0$  справедлива формула  $\langle e^{i\alpha} \rangle = \exp(-\bar{\alpha}^2/2)$ , находим

$$\langle v \rangle = v_0 = \langle e^{iS} \rangle = e^{-\sigma_S^2/2},$$

$$\Gamma_v(\rho) = \langle v(\rho_1) v^*(\rho_2) \rangle = \langle \exp \{ i [S(\rho_1) - S(\rho_2)] \} \rangle = \exp \left\{ -\frac{1}{2} \langle [S(\rho_1) - S(\rho_2)]^2 \rangle \right\}.$$

Но при  $\bar{S}=0$  средний квадрат разности фаз—это структурная функция фазы  $D_S(\rho)$ , связанная с корреляционной функцией  $\psi_S(\rho)$  соотношением (4.6). Поэтому

$$\Gamma_v(\rho) = \exp[-1/2 D_S(\rho)] = \exp[\psi_S(\rho) - \psi_S(0)] = \exp\{\sigma_S^2 [K_S(\rho) - 1]\}.$$

Среднее значение и поперечную функцию когерентности поля за экраном можно найти по формулам (10.3) и (10.14):

$$\bar{u} = v_0 e^{ikz} = \exp\{-\sigma_S^2/2 + ikz\},$$

$$\Gamma_{\perp}(\rho) = \Gamma_v(\rho) = \exp[-1/2 D_S(\rho)],$$

причем поперечная функция корреляции равна

$$\psi_{\perp}(\rho) = \psi_v(\rho) = \exp\{\sigma_S^2 [K_S(\rho) - 1]\} = e^{-\sigma_S^2}. \quad (10.15)$$

Если флуктуации фазы на экране слабы ( $\sigma_S^2 \ll 1$ ), то из (10.15) имеем

$$\psi_{\perp}(\rho) \approx \sigma_S^2 K_S(\rho) = \psi_S(\rho) \quad (\sigma_S^2 \ll 1), \quad (10.16)$$

т. е. при малых флуктуациях фазы поперечная функция корреляции поля во всем полупространстве  $z > 0$  совпадает с  $\psi_S(\rho)$ . В случае же больших дисперсий фазы ( $\sigma_S^2 \gg 1$ ) среднее значение поля  $v_0 = e^{-\sigma_S^2/2}$  пренебрежимо мало по сравнению с единицей, а величина  $\exp\{\sigma_S^2 [K_S(\rho) - 1]\}$  заметно отличается от нуля только при малых  $\rho$ . Учитывая это, пренебрежем в (10.15) членом  $e^{-\sigma_S^2}$  и разложим коэффициент корреляции  $K_S(\rho)$  в ряд Тейлора.

Считая для простоты флуктуации фазы на экране изотропными, имеем

$$K_S(\rho) = 1 + K_S''(0) \frac{\rho^2}{2} + \dots, \quad (10.17)$$

где штрихом обозначено дифференцирование по  $\rho$  (линейный по  $\rho$  член в разложении (10.17) отсутствует, так как случайное поле фазы  $S(\rho)$  предполагается дифференцируемым,  $K_S'(0) = 0$ ). В результате получаем

$$\psi_{\perp}(\rho) \approx e^{1/2 \sigma_S^2 K_S''(0) \rho^2} = e^{-1/2 |\psi_S''(0)| \rho^2} \quad (\sigma_S^2 \gg 1), \quad (10.18)$$

где учтено, что значения в нуле вторых производных  $K_S''(0)$  и  $I_S''(0) = \sigma_S^2 K_S''(0)$  отрицательны.

Из (10.18) видно, что корреляция исчезает при  $\rho \gg |\sigma_S^2 K_S''(0)|^{-1/2} \approx l_S/\sigma_S$ , где  $l_S \sim |K_S''(0)|^{-1/2}$  — радиус корреляции фазы. Поэтому для поперечного радиуса корреляции получаем оценку

$$l_{\perp} \sim l_S/\sigma_S \quad (\sigma_S \gg 1).$$

Таким образом, при  $\sigma_S \gg 1$  радиус корреляции примерно в  $\sigma_S$  раз меньше корреляционного масштаба фазы  $l_S$ . Нетрудно понять, с чем связано это различие масштабов  $l_{\perp}$  и  $l_S$ .

При смещении вдоль экрана на расстояние порядка радиуса корреляции  $l_S$  фаза  $S$  изменяется на величину  $\Delta S \sim \sigma_S$ , совершив при этом не более одной осцилляции. В то же время граничное поле  $v(\rho) = e^{iS}$  испытывает на том же расстоянии  $\rho \sim l_S$  примерно  $n \sim \sigma_S/\pi$  осцилляций, откуда и следует, что  $l_{\perp} \sim l_S/n \sim l_S/\sigma_S \ll l_S$ .

Приведенные выше соотношения используются, например, при интерпретации данных о прохождении ультракоротких радиоволн от внеземных источников через ионосферу Земли [23—26], в которой имеются неоднородности электронной концентрации. При определенных условиях можно считать, что поле  $v(\rho)$  на выходе из ионосферы испытывает только фазовые флуктуации.

Если дисперсия фазы  $\sigma_S^2$  мала, по сравнению с единицей, то пространственная функция корреляции поля на поверхности Земли совпадает, согласно (10.16), с функцией корреляции фазы  $\psi_S(\rho)$ , которая определенным образом связана с функцией корреляции ионосферных неоднородностей. Следовательно, при  $\sigma_S^2 \ll 1$  можно непосредственно измерить корреляционную функцию фазы волны, прошедшей через ионосферу, и судить о неоднородностях ионосферы.

Значительно сложнее интерпретировать данные наблюдений при  $\sigma_S^2 \gg 1$ . В этом случае функция корреляции изотропного поля  $\psi_{\perp}(\rho)$  связана с  $\psi_S(\rho)$  соотношением (10.18), при помощи которого можно оценить лишь величину  $|\psi_S''(0)| \sim \sigma_S^2/l_S^2$ . Дополнительные сведения о  $\sigma_S$  и  $l_S$  можно извлечь из данных об относительных флуктуациях интенсивности волны (см. ниже), но и с привлечением этих данных добиться однозначной интерпретации трудно.

Дело в том, что ионосферные неоднородности расположены не на фиксированной высоте, а распределены (причем неравномерно) на высотах от 100 до 400 и более километров над уровнем Земли. Кроме того, эти неоднородности имеют широкий диапазон горизонтальных масштабов (от  $l \leq 1$  до  $l \geq 500$  км), их действие часто маскируется сильно фокусирующими образованиями («ионосферными линзами») и т. д. Поэтому наблюдения флуктуаций поля на Земле позволяют оценивать только грубые

характеристики ионосферных неоднородностей: их горизонтальные масштабы, степень анизотропии (неоднородности, как правило, вытянуты вдоль линий магнитного поля Земли), вероятную высоту их расположения, а также среднюю скорость перемещения (дрейфа) неоднородностей [28, 29].

3. Флуктуации амплитуды и фазы за безграничным фазовым экраном. Во многих приложениях, в частности в задачах радиосвязи и радионавигации, представляют самостоятельный интерес статистические характеристики амплитуды и фазы волны. Для их вычисления кроме первой функции корреляции комплексного поля  $\psi_{\perp}(\rho) = \langle \tilde{u}(\rho_1, z) \tilde{u}^*(\rho_2, z) \rangle$ , которая в случае плоской волны и статистической однородности экрана не меняется при удалении от последнего, необходима и вторая корреляционная функция  $\tilde{\psi}_{\perp}(\rho) = \langle \tilde{u}(\rho_1, z) \tilde{u}(\rho_2, z) \rangle$ .

Закон преобразования второй функции корреляции при удалении от экрана можно получить при помощи френелевского приближения (9.17), которое применимо к полям именно с крупномасштабными неоднородностями на границе. Используя (9.17) и опуская для краткости аргумент  $\omega$ , находим

$$\tilde{\psi}_{\perp}(\rho, z) = \left( \frac{ke^{ikz}}{2\pi iz} \right)^2 \iint_{-\infty}^{\infty} \tilde{\psi}_v(\rho' - \rho'') \exp \left\{ i \frac{k}{2z} [(\rho_1 - \rho')^2 + (\rho_2 - \rho'')^2] \right\} d^2\rho' d^2\rho'', \quad (10.19)$$

где  $\tilde{\psi}_v(\rho) = \langle \tilde{v}(\rho_1) \tilde{v}(\rho_1 - \rho) \rangle$  — вторая корреляционная функция поля на экране. Введем новые переменные интегрирования  $\xi = \rho' - \rho''$  и  $\eta = (\rho' + \rho'')/2$ . Интеграл по  $\eta$  легко вычисляется, при этом зависимость от  $\rho_1 + \rho_2$  из (10.19) выпадает и остается лишь зависимость от  $\rho_1 - \rho_2$ . Переобозначив в оставшемся интеграле  $\xi$  через  $\rho'$  и  $\rho_1 - \rho_2$  через  $\rho$ , получаем

$$\tilde{\psi}_{\perp}(\rho, z) = \frac{ke^{2ikz}}{4\pi iz} \int_{-\infty}^{\infty} \tilde{\psi}_v(\rho') \exp \left[ \frac{ik}{4z} (\rho - \rho')^2 \right] d^2\rho'. \quad (10.20)$$

Таким образом, вторая корреляционная функция преобразуется почти так же, как и само поле: отличие от (9.17) заключается лишь в том, что в (10.20) входит *удвоенное* расстояние от экрана  $2z$ .

Последующие выкладки упростятся, если вместо флуктуационной части поля  $\tilde{u}$  ввести вспомогательную случайную величину  $a = \tilde{u}e^{-ikz}$ ,  $\bar{a} = 0$ , которая представляет собой комплексную амплитуду флуктуаций поля. Первая и вторая корреляционные функции этой величины преобразуются по формулам, аналогичным



(10.13) и (10.20):

$$\psi_a(\rho, z) = \psi_v(\rho), \quad (10.21)$$

$$\bar{\psi}_a(\rho, z) = \frac{k}{4\pi iz} \int_{-\infty}^{\infty} \bar{\psi}_v(\rho') \exp \left[ \frac{ik}{4z} (\rho - \rho')^2 \right] d^2 \rho'. \quad (10.22)$$

В отличие от (10.20), в (10.22) не входит множитель  $e^{2ikz}$ .

Кроме того, удобно нормировать средний квадрат поля на экране  $v(\rho)$  к единице:

$$\langle |v|^2 \rangle = \sigma_v^2 + |v_0|^2 = 1. \quad (10.23)$$

Такая нормировка отвечает непоглощающим и неотражающим экранам, поскольку падающая волна единичной интенсивности  $I_a = |e^{ikz}|^2 = 1$  порождает за экраном волну с той же средней интенсивностью:

$$\bar{I}_a = \langle |u|^2 \rangle = \langle |v|^2 \rangle = 1.$$

Это условие, очевидно, выполнено для чисто фазового экрана.

Свяжем теперь величину  $a$  с амплитудой  $A$  и фазой  $S$  волны  $u = Ae^{iS+ikz}$ , распространяющейся за экраном. Имеем

$$u(\rho, z) = A(\rho, z) \exp[ikz + iS(\rho, z)] = \bar{u} + \tilde{u} = (v_0 + a) e^{ikz},$$

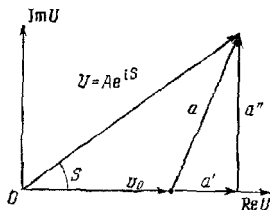


Рис. 3.

где  $v_0$  без ограничения общности можно считать вещественной величиной. Разделяя  $a$  на вещественную ( $a'$ ) и мнимую ( $a''$ ) части, находим, что

$$Ae^{iS} = v_0 + a = (v_0 + a') + ia'',$$

откуда

$$A = \sqrt{(v_0 + a')^2 + a''^2}, \quad S = \arctg \frac{a''}{v_0 + a'}. \quad (10.24)$$

На рис. 3 показаны соответствующие векторы на комплексной плоскости амплитуды  $U = Ae^{iS}$ .

Задача о нахождении из (10.24) статистических характеристик амплитуды  $A$  и фазы  $S$  решается до конца в двух частично перекрывающихся случаях: при *слабых флуктуациях* поля на экране, когда  $|a'|$  и  $|a''|$  малы по сравнению с единицей, и во *фраунгоферовой зоне*, когда величины  $a'$  и  $a''$  распределены по нормальному закону. Эти случаи и будут рассмотрены ниже.

Исследуем сначала флуктуации амплитуды и фазы при *слабых флуктуациях*. Пренебрегая малыми членами порядка  $|a'|^2$  и  $|a''|^2$ ,

и учитывая, что в силу (10.23)  $\sigma_0^2 + (a')^2 + (a'')^2 = 1$ , при помощи (10.24) получаем средние значения  $\bar{A}$  и  $\bar{S}$ :

$$\bar{A} = 1 - 1/2 \langle a'^2 \rangle, \quad \bar{S} = -\langle a' a'' \rangle, \quad (10.25)$$

а также выражения для корреляционных функций ( $\rho = \rho_1 - \rho_2$ ):

$$\begin{aligned} \psi_A(\rho, z) &\equiv \langle \bar{A}(\rho_1, z) \bar{A}(\rho_2, z) \rangle = \psi_{11}^A(\rho, z), \\ \psi_S(\rho, z) &\equiv \langle \bar{S}(\rho_1, z) \bar{S}(\rho_2, z) \rangle = \psi_{22}^S(\rho, z), \\ \psi_{AS}(\rho, z) &\equiv \langle \bar{A}(\rho_1, z) \bar{S}(\rho_2, z) \rangle = \psi_{12}^A(\rho, z), \end{aligned} \quad (10.26)$$

где  $\psi_{ij}^a(\rho, z)$ ,  $i, j = 1, 2$  — корреляционные функции вещественной и мнимой частей комплексного случайного поля  $a$ . В общем случае они выражаются через первую и вторую функции корреляции  $\psi_a(\rho, z)$  и  $\tilde{\psi}_a(\rho, z)$  при помощи формул (2.14). Однако, если  $\psi_a(\rho, z)$  — четная функция  $\rho$ , то справедливы более простые формулы (2.22):

$$\begin{aligned} \psi_{11}^a &= 1/2 (\psi_a + \text{Re } \tilde{\psi}_a), \quad \psi_{22}^a = 1/2 (\psi_a - \text{Re } \tilde{\psi}_a), \\ \psi_{12}^a &= \psi_{21}^a = 1/2 \text{Im } \tilde{\psi}_a. \end{aligned} \quad (10.27)$$

Таким образом, посредством формул (10.25) и (10.26) первые два момента амплитуды и фазы выражены через первую и вторую корреляционные функции комплексного поля  $a$ . Последние же преобразуются при удалении от экрана в соответствии с выражениями (10.21) и (10.22), чем и решается поставленная задача [23].

В частном случае фазового экрана с начальной фазой  $S$ , распределенной по нормальному закону, имеем

$$\psi_{A, S}(\rho, z) = \frac{1}{2} \left[ \psi_S^0(\rho) \mp \frac{k}{4\pi z} \int_{-\infty}^{\infty} \psi_S^0(\rho') \sin \frac{k(\rho - \rho')^2}{4z} d^2 \rho' \right], \quad (10.28)$$

$$\psi_{AS}(\rho, z) = \frac{k}{8\pi z} \int_{-\infty}^{\infty} \psi_S^0(\rho') \cos \frac{k(\rho - \rho')^2}{4z} d^2 \rho', \quad (10.29)$$

где  $\psi_S^0(\rho)$  — корреляционная функция фазы в плоскости  $z=0$ .

Для гауссовой корреляционной функции фазы

$$\psi_S^0(\rho) = \sigma_{S_0}^2 e^{-\rho^2/2l_S^2} \quad (10.30)$$

эти интегралы легко вычисляются и, в частности, при  $\rho=0$  дают

$$\sigma_{A, S}^2(z) = \frac{\sigma_{S_0}^2}{2} \left( 1 \mp \frac{1}{1 + D^2} \right), \quad \langle \tilde{A} \tilde{S} \rangle = -\frac{\sigma_{S_0}^2}{2} \frac{D}{1 + D^2}, \quad (10.31)$$

где  $D = D(z) = 2z/kl_S^2$ . Величину  $D$  называют вслед за Г. С. Гореликом [43] *волновым параметром*. Этот параметр показывает

во сколько раз площадь первой зоны Френеля  $\pi(\sqrt{\lambda z})^2 = 2\pi^2 z/k$  превышает по порядку величины «площадь» одной неоднородности  $\sim \pi l_0^2$ , т. е. сколько неоднородностей уместится в этой зоне. В зависимости от значения волнового параметра можно выделить три области дистанции  $z$  (которые тоже называют зонами): ближнюю ( $D \ll 1$ ), френелевскую ( $D \sim 1$ ) и фраунгоферову ( $D \gg 1$ ) зоны (по отношению к отдельной неоднородности). Для каждой из них характерны определенные особенности флуктуаций.

В ближней зоне ( $D \ll 1$ ) преобладают, естественно, фазовые флуктуации:  $\sigma_A^2 \ll \sigma_S^2 \approx \sigma_{S_0}^2$ . При удалении от экрана амплитудные флуктуации нарастают, а фазовые уменьшаются, причем в пределе  $D \rightarrow \infty$  (фраунгоферова зона) дисперсии амплитуды и фазы выравниваются:

$$\sigma_{A,S}^2|_{z \rightarrow \infty} = 1/2 \sigma_{S_0}^2. \quad (10.32)$$

Корреляция между  $A$  и  $S$  пренебрежима в ближней и дальней зонах и максимальна при  $D \sim 1$ .

Обратимся теперь к флуктуациям в зоне Фраунгофера ( $D \gg 1$ ) при произвольных (не обязательно слабых) флуктуациях поля на экране. При  $D \gg 1$  в пределах первой зоны Френеля с радиусом  $\sqrt{\lambda z}$ , которая только и существенна для интегрирования в (9.17), уместится много неоднородностей поля на экране. В силу центральной предельной теоремы теории вероятностей закон распределения величин  $a'$  и  $a''$  приближается поэтому к нормальному.

Нормализация этих величин обусловлена «фильтрующим» действием свободного пространства и имеет такую же природу, как и нормализация временных сигналов на выходе узкополосных фильтров (ч. I, § 50). Действительно, преобразование случайного поля по формуле (9.17) вполне аналогично преобразованию случайных процессов, причем аналогом импульсной функции в нашем случае является разностное ядро преобразования (9.17), которое и осуществляет фильтрацию с эффективной шириной полосы пространственных частот  $\Delta k \sim \sqrt{k/z} \sim (\lambda z)^{-1/2}$ . С ростом дистанции  $z$  эта полоса сужается и при  $z \gg kl_0^2$  (т. е. в дальней зоне,  $D \gg 1$ ) становится значительно уже первоначальной ширины пространственного спектра  $k_g \sim 1/l_0$ . При этих условиях и происходит нормализация поля  $u(\rho, z)$ .

То обстоятельство, что  $a'$  и  $a''$  в дальней зоне распределены по нормальному закону, дает возможность найти плотности вероятностей амплитуды и фазы и вычислить моменты этих величин. По существу, речь идет о хорошо изученной задаче о статистике огибающей  $A$  и фазы  $S$  сигнала  $u(\rho, z) = Ae^{iS+ikz} = (v_0 + a)e^{ikz}$ , представляющего собой сумму гармонического колебания  $v_0 e^{ikz}$  и гауссова шума  $\tilde{u} = ae^{ikz}$ .

Для гауссовых величин закон распределения вероятностей полностью характеризуется только низшими моментами — средними значениями и функциями корреляции. В нашем случае  $\bar{a}' = \bar{a}'' = 0$ , а функции корреляции  $\psi_{11}^a$ ,  $\psi_{12}^a$  и  $\psi_{22}^a$  выражаются через первую и вторую функции корреляции комплексного поля  $a$  посредством формул (10.27). В зоне Фраунгофера эти формулы существенно упрощаются, поскольку при  $D \gg 1$  вторая корреляционная функция  $\bar{\psi}_a(\rho, z)$  пренебрежимо мала по сравнению с первой. В результате при  $D \gg 1$  имеем

$$\psi_{11}^a(\rho, z) = \psi_{22}^a(\rho, z) = 1/2 \psi_v(\rho), \quad \psi_{12}^a(\rho, z) = 0, \quad (10.33)$$

т. е. в зоне Фраунгофера поля  $a'$  и  $a''$  некоррелированы, а их автокорреляционные функции одинаковы и отличаются коэффициентом  $1/2$  от первой функции корреляции поля на экране. Как следствие этого, функции  $\psi_{11}^a$  и  $\psi_{22}^a$  и связанные с ними статистические характеристики амплитуды и фазы не зависят от расстояния до экрана: с ростом дистанции  $z$  они остаются такими же, как на «входе» в дальнюю зону  $z \gg kl_0^2$ .

Определение статистических моментов амплитуды и фазы при известной (гауссовой) статистике  $a'$  и  $a''$  проводится таким же образом, как и в ч. I, §§ 25 и 44 (см. также [30, 31]).

4. Флуктуации интенсивности за безграничным фазовым экраном. При рассмотрении флуктуаций интенсивности обычно интересуются их функцией корреляции ( $\rho = \rho_1 - \rho_2$ )

$$\psi_I(\rho, z) = \langle I(\rho_1, z) \bar{I}(\rho_2, z) \rangle = \langle I(\rho_1, z) I(\rho_2, z) \rangle - \bar{I}^2 \quad (10.34)$$

и так называемым индексом мерцаний  $\beta$ :

$$\beta^2(z) = \frac{\langle I^2(z) \rangle - \bar{I}^2(z)}{\bar{I}^2} = \frac{\sigma_I^2(z)}{\bar{I}^2}, \quad (10.35)$$

который характеризует относительные флуктуации интенсивности. Если принять, как и выше, что  $\bar{I} = 1$ , то

$$\psi_I(\rho, z) = \langle I(\rho_1, z) I(\rho_2, z) \rangle - 1, \quad (10.34')$$

$$\beta^2(z) = \langle I^2(z) \rangle - 1 = \sigma_I^2(z). \quad (10.35')$$

Как и при анализе амплитудных и фазовых флуктуаций, расчеты  $\psi_I$  и  $\beta$  удастся довести до конца лишь в двух частично пересекающихся предельных случаях — для слабых флуктуаций и для фразунгоферовой зоны.

Используя результаты п. 3, легко показать, что в случае слабых флуктуаций

$$\psi_I(\rho, z) = 4\psi_A(\rho, z), \\ \beta^2(z) = 4\sigma_A^2(z).$$

Таким образом, флуктуации интенсивности меняются при удалении от экрана по тому же закону, что и флуктуации амплитуды. При  $D \gg 1$  (зона Фраунгофера)  $\beta^2(z)$  стремится к предельному значению

$$\beta_\infty^2 = 2\sigma_{S_0}^2, \quad \sigma_{S_0}^2 \ll 1, \quad (10.36)$$

независимо от вида функции корреляции фазы на экране.

При расчете  $\psi_I$  в *дальней зоне* ( $D \gg 1$ ) воспользуемся формулой из задачи 12 к гл. 1, которая справедлива для полей, распределенных по нормальному закону. При  $D \gg 1$  имеем  $\bar{\psi}_u = e^{i k z} \bar{\psi}_u \approx 0$ , так что  $\bar{\Gamma}(1, 2) = \bar{u}(1) \bar{u}(2)$  и

$$\psi_I(\rho, z) = |\Gamma_u(\rho)|^2 - |\bar{u}|^4 = |\Gamma_v(\rho)|^2 - |v_0|^4. \quad (10.37)$$

Так как мы приняли, что  $\Gamma_v(0) = \bar{I} = 1$ , для индекса мерцаний в дальней зоне ( $z \rightarrow \infty$ ) получаем

$$\beta_z^2 \rightarrow \infty \equiv \beta_\infty^2 = \psi_I(0) = 1 - |v_0|^4. \quad (10.38)$$

Согласно этим формулам функция корреляции интенсивности  $\psi_I$  и индекс мерцаний  $\beta^2$  при  $D \gg 1$  уже не зависят от  $z$ .

В частном случае фазового экрана с гауссовыми флуктуациями фазы, когда  $v_0 = e^{-\sigma_{S_0}^2/2}$ , а  $\psi_v(\rho)$  дается выражением (10.15), получаем

$$\psi_I(\rho) = e^{-2\sigma_{S_0}^2} \{ \exp[2\psi_S^2(\rho)] - 1 \}, \quad \beta^2 = 1 - e^{-2\sigma_{S_0}^2}. \quad (10.39)$$

При  $\sigma_{S_0}^2 \ll 1$  последнее выражение переходит, как и следовало ожидать, в формулу (10.36). В другом частном случае — экран с нулевым средним полем ( $v_0 = 0$ ) — выражения (10.37) и (10.38) принимают особенно простой вид:

$$\psi_I(\rho) = |\Gamma_v(\rho)|^2 = |\psi_v(\rho)|^2, \quad \beta_\infty^2 = 1.$$

Анализ флуктуаций в общем случае наталкивается на значительные математические трудности. Если флуктуации поля на экране не малы и точка наблюдения не удалена во фраунгоферову зону по отношению к отдельной неоднородности, то расчет индекса мерцаний сводится, в рамках френелевского приближения (9.17), к вычислению интеграла

$$\beta^2(z) = \left( \frac{k}{2\pi z} \right)^4 \iiint_{-\infty}^{\infty} \Gamma_v^{(2,2)}(1, 2, 3, 4) \times \\ \times \exp \left[ \frac{ik}{2z} (\rho_1^2 + \rho_2^2 - \rho_3^2 - \rho_4^2) \right] d^2\rho_1 \dots d^2\rho_4 - 1, \quad (10.40)$$

где

$$\Gamma_{\psi}^{(2,2)}(1, 2, 3, 4) = \langle \psi(\rho_1) \psi(\rho_2) \psi^*(\rho_3) \psi^*(\rho_4) \rangle$$

— смешанный момент четвертого порядка. Для произвольных функций  $\Gamma_{\psi}^{(2,2)}(1, 2, 3, 4)$  значения интегралов вида (10.40) можно находить лишь численными методами. Рассмотрим полученные таким путем в [32] результаты для фазового экрана.

Если фазовые флуктуации распределены по нормальному закону, то

$$\Gamma_{\psi}^{(2,2)}(1, 2, 3, 4) = \langle \exp \{ i [S(\rho_1) + S(\rho_2) - S(\rho_3) - S(\rho_4)] \} \rangle = \\ = \exp \{ - [2\sigma_{S_0}^2 - \psi_{13} - \psi_{14} - \psi_{23} - \psi_{24} + \psi_{12} + \psi_{34}] \},$$

где  $\psi_{ik} = \psi_S(\rho_i - \rho_k)$  — значения функции корреляции фазы  $S$  при  $\rho = \rho_i - \rho_k$ , а  $\sigma_{S_0}^2$  — дисперсия фазы. На рис. 4 показаны полученные в [32] графики зависимости индекса мерцаний  $\beta^2$  от волнового параметра  $D = 2z/kl_S^2$  для гауссовой функции корреляции фазы (10.30) и для значений дисперсии  $\sigma_{S_0}^2$  от 0,1 до 5.

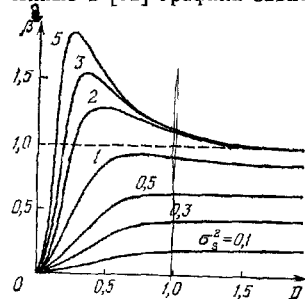


Рис. 4.

Поведение этих кривых нам уже частично известно: в зоне Фраунгофера ( $D \gg 1$ )  $\beta^2 = \beta_\infty^2 = 1 - e^{-2\sigma_{S_0}^2}$ , а при слабых флуктуациях фазы ( $\sigma_{S_0}^2 \ll 1$ )  $\beta^2 = 2\sigma_{S_0}^2 D^2 / (1 + D^2)$ . Новым является наличие при  $\sigma_{S_0}^2 > 1$  максимумов  $\beta^2$  при промежуточных значениях  $D \ll 1$ . Для величины максимумов  $\beta^2$  при сильных флуктуациях фазы ( $\sigma_{S_0}^2 \gg 1$ ) в [32], а также в [33], где было проведено качественное рассмотрение вопроса, дана оценка  $\beta_{\max}^2 \sim \ln \sigma_{S_0}^2$ . Более аккуратные расчеты, проделанные в [34, 35] для функции корреляции фазы  $\psi_S(\rho)$  произвольного вида, привели к следующему асимптотическому значению  $\beta_{\max}^2$  при  $\sigma_{S_0}^2 \gg 1$ :

$$\beta_{\max}^2 \approx \frac{1}{\pi e} \ln \left( \frac{2\sigma_{S_0}^2}{\gamma} \right).$$

Здесь  $\gamma$  — численный коэффициент, равный  $\frac{l_1^4}{6} \frac{\partial^4 K_S(0)}{\partial \rho^4}$ , а  $l_1 = |K_S^*(0)|^{-1/2}$  — величина порядка радиуса корреляции фазы. Для гауссовой функции корреляции (10.30)  $\gamma = 1/2$ ,  $l_1 = l_S$ . Эти оценки указывают на логарифмическое «насыщение» индекса мерцаний  $\beta_{\max}^2$  с ростом  $\sigma_{S_0}^2$ .

Максимум индекса мерцаний приходится, согласно [32—35], на расстояние  $z_m$ , равное

$$z_m = \frac{kl_S^2}{\sigma_{S0}}. \quad (10.41)$$

Формула (10.41), как и самый факт появления максимумов, допускает простую геометрическую интерпретацию: флуктуации интенсивности максимальны там, где волны за фазовым экраном *фокусируются*. Действительно, в приближении геометрической оптики фокусировка происходит на расстоянии  $z_m \sim 1/\nu$  от фазового экрана, где  $\nu$  — кривизна фазового фронта. По порядку величины  $\nu \sim \frac{1}{k} \left( \frac{\partial^2 S}{\partial x^2} \right)$  ( $x$  — одна из координат в плоскости  $z=0$ ) и  $\frac{\partial^2 S}{\partial x^2} \sim \frac{\sigma_{S0}}{l_S^2}$ . Следовательно,  $z_m \sim 1/\nu \sim kl_S^2/\sigma_{S0}$ , что совпадает с (10.41). Очевидно, чем больше дисперсия фазы  $\sigma_{S0}^2$ , тем ближе к экрану расположена зона фокусировок.

Эффект фокусировки волн и обусловленные им максимумы индекса мерцаний характерны, очевидно, только при условии, что на экране хорошо выражены именно фазовые флуктуации. Если же на экране флуктуирует только амплитуда, то дифракция приводит не к увеличению, а, наоборот, к *сглаживанию* флуктуаций интенсивности, т. е. к уменьшению индекса мерцаний. Примеры такого сглаживания рассмотрены в задачах 7 и 8. Сглаживание флуктуаций интенсивности происходит также в том случае, когда на фазовый экран падает не плоская волна, а волна от источника конечных угловых размеров [32].

## § 11. Дифракция случайных полей в простейших оптических системах

В предыдущем параграфе мы рассмотрели одну из простейших статистических дифракционных задач — дифракцию *детерминированной* плоской и монохроматической волны на бесконечном и статистически однородном *хаотическом* экране. Здесь мы обратимся к дифракции *случайных* полей на *детерминированных* объектах. В основном мы будем иметь в виду слабо расходящиеся волновые пучки, чаще всего встречающиеся в оптических и квазиоптических системах.

1. Прохождение случайной волны через отверстие в экране. Теорема Ван-Циттерта—Цернике. Пусть  $u(\omega, \rho) = u_n(\omega, \mathbf{r})|_{z=0}$  — спектральная амплитуда поля, создаваемого каким-либо источником в плоскости  $z=0$ . Если поместить в этой плоскости непрозрачный экран с отверстием  $S$ , вырезающим пучок волн конечного диаметра, то непосредственно за экраном образуется поле «на выходе»  $u_v(\omega, \rho)$ , которое в

приближении Кирхгофа равно<sup>1)</sup>

$$u_v(\omega, \rho) = \begin{cases} v(\omega, \rho) & \text{на } S, \\ 0 & \text{вне } S. \end{cases}$$

Это поле можно записать через функцию пропускания отверстия:

$$u_v(\omega, \rho) = M(\rho) v(\omega, \rho), \quad (11.1)$$

считая, что

$$M(\rho) = \begin{cases} 1 & \text{на } S, \\ 0 & \text{вне } S. \end{cases} \quad (11.2)$$

В случае плавных (в масштабе длины волны) флуктуаций граничного поля  $v(\omega, \rho)$  для вычисления поля за отверстием можно воспользоваться френелевским приближением (9.17), которое с учетом (11.1) дает

$$u(\omega, \mathbf{r}) = \frac{ke^{ikz}}{2\pi iz} \int_{-\infty}^{\infty} M(\rho') v(\omega, \rho') \exp \left[ \frac{ik}{2z} (\rho - \rho')^2 \right] d^2\rho'. \quad (11.3)$$

Среднее по ансамблю источников значение  $v(\omega, \rho)$  равно нулю, вследствие чего и

$$\langle u(\omega, \mathbf{r}) \rangle = 0. \quad (11.4)$$

Вычислим *поперечную* функцию когерентности дифракционного поля, т. е. функцию когерентности в плоскости  $z = \text{const}$ . В соответствии с (11.3) имеем

$$\begin{aligned} \Gamma_{\perp}(\omega, \rho_1, \rho_2, z) &= \langle u(\omega, \rho_1, z) u^*(\omega, \rho_2, z) \rangle = \\ &= \left( \frac{k}{2\pi z} \right)^2 \iint_{-\infty}^{\infty} M(\rho') M(\rho'') \Gamma_v(\omega, \rho' - \rho'') \times \\ &\times \exp \left\{ \frac{ik}{2z} [(\rho_1 - \rho')^2 - (\rho_2 - \rho'')^2] \right\} d^2\rho' d^2\rho'', \end{aligned} \quad (11.5)$$

где  $\Gamma_v(\omega, \rho' - \rho'')$  — пространственная функция когерентности граничного поля  $v(\omega, \rho)$ , которое предполагается статистически однородным. Пространственно-временные функции когерентности полей  $u(t, \mathbf{r})$  и  $v(t, \rho)$  связаны с  $\Gamma_{\perp}(\omega, \rho_1, \rho_2, z)$  и  $\Gamma_v(\omega, \rho' - \rho'')$  преобразованиями Фурье, например:

$$\Gamma_{\perp}(\tau, \rho_1, \rho_2, z) = \int_0^{\infty} \Gamma_{\perp}(\omega, \rho_1, \rho_2, z) e^{-i\omega\tau} d\omega. \quad (11.6)$$

<sup>1)</sup> Приближение Кирхгофа применимо, как известно, при условии, что размеры отверстия  $a$  велики по сравнению с длиной волны,  $a \gg \lambda$ .



В силу (11.4) все эти функции когерентности совпадают с соответствующими функциями корреляции.

Не конкретизируя вида функции когерентности (корреляции) граничного поля, интеграл (11.5) можно вычислить в двух предельных случаях — при больших и малых размерах отверстия  $a$  по сравнению с радиусом корреляции  $l_v$  (но всегда при  $a \gg \lambda$ ).

При  $a \ll l_v$  (малое отверстие) функция когерентности граничного поля  $\Gamma_v$  практически постоянна в пределах отверстия и ее можно вынести за знак интеграла со значением  $\Gamma_v(\omega, 0) \equiv \bar{\Gamma}_v(\omega)$ . В результате<sup>1)</sup>

$$\Gamma_{\perp}(\rho_1, \rho_2, z) = \bar{\Gamma}_v \mu_M(\rho_1, z) \mu_M^*(\rho_2, z), \quad (11.7)$$

где поле  $\mu_M(\rho, z)$  дается выражением

$$\mu_M(\rho, z) = \frac{ke^{ikz}}{2\pi iz} \int M(\rho') \exp \left[ \frac{ik}{2z} (\rho - \rho')^2 \right] d^2 \rho' \quad (11.8)$$

и представляет собой дифракционное поле за отверстием при нормальном падении на него плоской волны единичной амплитуды.

Коэффициент корреляции при дифракции на малом отверстии равен

$$K_{\perp}(\rho_1, \rho_2, z) = \frac{\Gamma_{\perp}(\rho_1, \rho_2, z)}{\sqrt{\Gamma_{\perp}(\rho_1, \rho_1, z) \Gamma_{\perp}(\rho_2, \rho_2, z)}} = \frac{\mu_M(\rho_1, z) \mu_M^*(\rho_2, z)}{|\mu_M(\rho_1, z)| |\mu_M(\rho_2, z)|},$$

так что  $|K_{\perp}| = 1$ . Это означает, что случайная волна, прошедшая через малое отверстие, порождает *пространственно когерентное* поле. До создания лазеров пропускание света через малое отверстие (наряду с использованием малых источников света) было практически единственным способом получения пространственно когерентного света. Этим способом пользуются и в настоящее время, если не требуется высокой интенсивности поля.

Обратимся к другому предельному случаю  $a \gg l_v$  (широкое отверстие). Прежде всего отметим, что вблизи отверстия имеется область, в которой функция корреляции будет такой же, как и в отсутствие непрозрачного экрана. Форму и размеры этой области можно оценить, используя спектральное представление. Ширина двумерного спектра  $F_v(\mathbf{x})$  случайного поля  $v$  имеет порядок  $x_v \sim 1/l_v$  или, в пересчете на углы,  $\delta \sim x_v/k \sim \lambda/l_v$ . Следовательно, если точки наблюдения  $\mathbf{r}_1$  и  $\mathbf{r}_2$  находятся внутри

<sup>1)</sup> Начиная с формулы (11.7), мы опускаем аргумент  $\omega$  всюду, где это не может привести к недоразумениям.

конуса с основанием  $a$  и углом при вершине  $\delta$  (рис. 5), то поле «не почувствует» влияния краев диафрагмы. Для всех точек внутри этого конуса поперечная функция корреляции будет такой же, как у граничного поля:

$$\Gamma_{\perp}(\rho_1, \rho_2, z) = \Gamma_v(\rho_1 - \rho_2).$$

Предельная дистанция  $z_{\text{пр}}$ , на которой еще справедливо это соотношение, оценивается как <sup>1)</sup>

$$z_{\text{пр}} \sim \frac{a}{\delta} \sim \frac{al_v}{\lambda} \sim kal_v. \quad (11.9)$$

При  $a \gg l_v$  эта дистанция отвечает дальней зоне по отношению к размеру неоднородностей  $l_v$ , так как

$$\frac{z_{\text{пр}}}{kl_v^2} \sim \frac{a}{l_v} \gg 1,$$

и ближней зоне по отношению к размеру отверстия  $a$ :

$$\frac{z_{\text{пр}}}{ka^2} \sim \frac{l_v}{a} \ll 1.$$

Отсюда, в частности, следует практически важный вывод, что распределение флуктуаций поля за отверстием нормализуется еще в ближней зоне апертуры  $a$ .

Вычислим интеграл (11.5) при  $z \gg z_{\text{пр}}$ . Для этого перейдем к новым переменным  $\xi = \rho' - \rho''$  и  $\eta = (\rho' + \rho'')/2$  и обозначим полусумму  $(\rho_1 + \rho_2)/2$  через  $\rho_+$ , а разность  $\rho_1 - \rho_2$  — через  $\rho$ . Тогда (11.5) принимает вид

$$\Gamma_{\perp}(\rho, \rho_+, z) =$$

$$= \left(\frac{k}{2\pi z}\right)^2 \exp\left(\frac{ik\rho\rho_+}{z}\right) \iint_{-\infty}^{\infty} M\left(\eta + \frac{\xi}{2}\right) M\left(\eta - \frac{\xi}{2}\right) \Gamma_v(\xi) \times \\ \times \exp\left\{\frac{ik}{z}[\xi\eta - \eta\rho - \xi\rho_+]\right\} d^2\xi d^2\eta.$$

Область интегрирования по  $\xi$  ограничена здесь неравенством  $\xi \ll l_a \ll a$ , так как при  $\xi \gg l_v$  функция корреляции  $\Gamma_v(\xi)$  близка к нулю. В то же время  $\eta \ll a$ , поскольку при  $\eta > a$  обращается в нуль произведение  $M(\eta + \xi/2) M(\eta - \xi/2) \approx M^2(\eta) = M(\eta)$ . Поэтому первое слагаемое в показателе экспоненты ограничено сверху неравенством  $k\xi\eta/z \ll kal_v/z \sim z_{\text{пр}}/z$ . При  $z \gg z_{\text{пр}}$  этим слагаемым

<sup>1)</sup> Если  $l_v \ll a$ , то в качестве  $z_{\text{пр}}$  следует брать величину  $z_{\text{пр}} \sim a$ , так как в этом случае ширина углового спектра  $\delta$  сравнима с  $\pi/2$ .

можно пренебречь, и тогда интегралы по  $\xi$  и  $\eta$  разделяются

$$\Gamma_{\perp}(\rho, \rho_+, z) = \left(\frac{k}{2\pi z}\right)^2 \exp\left(\frac{ik\rho\rho_+}{z}\right) \int_{-\infty}^{\infty} d^2\eta M(\eta) \exp\left(-\frac{ik\eta\rho}{z}\right) \times \\ \times \int_{-\infty}^{\infty} d^2\xi \Gamma_v(\xi) \exp\left(-\frac{ik\xi\rho_+}{z}\right). \quad (11.10)$$

Точно такое же выражение можно получить и непосредственно из формулы (9.18) для поля во френгоферовой зоне:

$$u(\rho, z) = \frac{k \exp\left[ik\left(z + \frac{\rho^2}{2z}\right)\right]}{2\pi iz} \int_{-\infty}^{\infty} M(\rho') v(\rho') \exp\left(-\frac{ik\rho\rho'}{z}\right) d^2\rho'. \quad (11.11)$$

Однако приближение (11.11) справедливо только в дальней зоне  $z \gg ka^2$ , а формула (11.10) применима и на гораздо меньших дистанциях  $z \gg z_{\text{фр}} \sim ka l_v$ . Иными словами, функция когерентности (11.10) формируется еще в ближней зоне по отношению к отверстию.

Зависимость функции когерентности (11.10) от поперечных координат  $\rho = \rho_1 - \rho_2$  и  $\rho_+ = (\rho_1 + \rho_2)/2$  определяется произведением трех сомножителей. Множитель перед интегралом

$$\exp\{ik\rho\rho_+/z\} = \exp\{ik(\rho_1^2 - \rho_2^2)/2z\} \approx \exp\{ik(r_1 - r_2)\}$$

возникает просто из-за того, что сферические волны, покидающие отверстие, не согласованы с плоской формой поверхности  $z = \text{const}$ , на которой рассматривается корреляция поля. Если взять корреляцию на сфере  $r = \text{const}$ , этот множитель обратится в единицу.

Интеграл по  $\eta$  в (11.10), зависящий только от разности  $\rho = \rho_1 - \rho_2$ , характеризует поле когерентного источника с распределением амплитуды  $M(\eta)$ . При размере отверстия  $a$  угловая ширина диаграммы направленности такого излучателя имеет порядок  $\lambda/a$ , а линейная ширина центрального дифракционного лепестка в плоскости  $z = \text{const}$  равна  $\Delta\rho \sim \lambda z/a$ . Отсюда для поперечного радиуса корреляции сразу же следует, что

$$l_{\perp} \sim \frac{\lambda z}{a} \sim \frac{\lambda}{\gamma}, \quad (11.12)$$

где  $\gamma \sim a/z$  — угловой размер отверстия из точки наблюдения. Как и ширина дифракционных лепестков, поперечный радиус корреляции возрастает по мере удаления от плоскости  $z = 0$ . Увеличение радиуса когерентности геометрически объясняется

тем, что при удалении от отверстия фазовые фронты всех элементарных сферических волн можно считать (с точностью до  $\lambda$ ) совпадающими на все большей площади.

Второй интеграл в (11.10) определяет пространственное распределение средней интенсивности дифракционного поля. Положив в (11.10)  $\rho=0$  и учитывая, что интеграл по  $\eta$  при  $\rho=0$  равен площади отверстия  $S$ , имеем

$$\bar{I}_u(\rho_+, z) = \frac{k^2 S}{(2\pi z)^2} \int_{-\infty}^{\infty} \Gamma_v(\xi) \exp\left(-\frac{ik\rho_+ \xi}{z}\right) d^2 \xi. \quad (11.13)$$

По предположению масштаб  $l_v$  изменения функции  $\Gamma_v$  мал по сравнению с поперечником отверстия  $a$ . Поэтому ширина распределения интенсивности  $\bar{I}_u$  в плоскости  $z = \text{const}$ , составляющая  $\rho_+ \sim \lambda z / l_v$  (угловая ширина имеет порядок  $\rho_+ / z \sim \lambda / l_v \sim 1 / kl_v$ ),

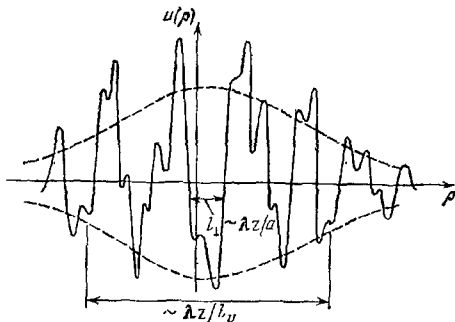


Рис. 6.

велика по сравнению с поперечным радиусом корреляции  $l_{\perp}$ :

$$\frac{\lambda z}{l_v} \gg l_{\perp} \sim \frac{\lambda z}{a} \quad (l_v \ll a).$$

Таким образом, мы имеем здесь дело с примером квазиоднородного (в плоскости  $z = \text{const}$ ) поля. В данном случае все реализации дифракционного поля за отверстием представляют собой быстрые пространственные осцилляции с масштабом  $l_{\perp}$  и с размахом порядка  $\sqrt{\bar{I}_u(\rho)}$  (рис. 6.).

Если разделить функцию когерентности (11.10) на среднюю интенсивность дифракционного поля, то мы получим коэффи-

коэффициент поперечной пространственной корреляции поля

$$K_{\perp}(\rho) = \exp\left(\frac{ik\rho\rho_+}{z}\right) \mathcal{F}_M\left(\frac{k\rho}{z}\right),$$

где величина

$$\mathcal{F}_M(\kappa) = \frac{1}{S} \int_{-\infty}^{\infty} M(\eta) \exp(-i\kappa\eta) d^2\eta \quad (11.14)$$

представляет собой нормированное к единице ( $\mathcal{F}(0) = 1$ ) преобразование Фурье от  $M(\eta)$  и одновременно — диаграмму направленности отверстия при нормальном падении на него плоской монохроматической волны. В частном случае круглого отверстия, когда  $M(\rho) = 1$  при  $\rho < a$  и обращается в нуль вне этого круга,

$$\mathcal{F}_M(\kappa) = 2J_1(\kappa a)/\kappa a,$$

где  $J_1(\kappa a)$  — функция Бесселя первого порядка.

В пределе, когда радиус корреляции граничного поля  $l_v$  мал по сравнению с длиной волны, интеграл по  $\xi$  перестает зависеть от  $\rho_+$ :

$$\int_{-\infty}^{\infty} \Gamma_v(\xi) \exp\left(-\frac{ik\rho+\xi}{z}\right) d^2\xi \approx \bar{I}_v S_v \quad (l_v \ll \lambda).$$

Через  $S_v$  здесь обозначена эффективная площадь когерентности граничного поля:

$$S_v = \frac{1}{\bar{I}_v} \int_{-\infty}^{\infty} \Gamma_v(\xi) d^2\xi = \int_{-\infty}^{\infty} K_v(\xi) d^2\xi, \quad (11.15)$$

которая по порядку величины равна  $l_v^2$ . В результате при  $l_v \ll \lambda$  из (11.10) вытекает так называемая *теорема Ван-Циттерта — Цернике*:

$$\Gamma_v(\rho, \rho_+, z) = \frac{1}{z^2} \exp\left(\frac{ik\rho\rho_+}{z}\right) \int_{-\infty}^{\infty} J(\rho') \exp\left(-\frac{ik\rho\rho'}{z}\right) d^2\rho'. \quad (11.16)$$

Входящую сюда величину

$$J(\rho') = \frac{k^2}{4\pi^2} \bar{I}_v S_v M(\rho') \quad (11.17)$$

называют *приведенной интенсивностью*.

Согласно теореме Ван-Циттерта — Цернике модуль поперечной функции когерентности  $\Gamma_{\perp}$  зависит от разности  $\rho = \rho_1 - \rho_2$  так

же, как и поле в зоне Фраунгофера, создаваемое полностью когерентным источником с распределением амплитуды, пропорциональным  $J(\rho')$ . От обычных дифракционных формул выражение (11.16) отличается в двух отношениях. Во-первых, квадратичная по полю величина  $\Gamma_{\perp}$  обратно пропорциональна квадрату (а не первой степени) расстояния. Во-вторых, это выражение вместе с исходной формулой (11.10) применимо, как уже отмечалось, на дистанциях  $z \gg z_{\text{нр}}$ , причем в рассматриваемом пределе  $l_v \ll a$  величина  $z_{\text{нр}}$  сравнима с  $a$ , тогда как диаграмма когерентного источника с поперечником  $a$  формируется на значительно больших дистанциях  $z \gg ka^2 \gg z_{\text{нр}} \sim a$ .

Обычно теорему Ван-Циттерта—Цернике (11.16) получают путем формального введения дельта-коррелированных флуктуаций граничного поля, полагая в (11.5) или в (11.10)

$$\Gamma_v(\rho - \rho') = \bar{I}_v(\rho') \delta(\rho' - \rho'') S_v, \quad (11.18)$$

при этом в выражении (11.16) для приведенной интенсивности вместо произведения  $\bar{I}_v M(\rho')$  возникает переменная интенсивность  $\bar{I}_v(\rho')$ . Источник с независимыми значениями поля в сколь угодно близких точках называют пространственно *некогерентным*.

Представление о пространственно некогерентном источнике является идеализацией, имеющей ограниченную область применимости. Дело в том, что поле *бегущих* волн по самой своей природе не может быть дельта-коррелированным в пространстве, поскольку (§ 9) масштаб изменения поля  $l_v$  не может быть меньше длины волны  $\lambda^1$ ). Поэтому в рассматриваемой постановке задачи (дифракция случайного поля на отверстии) переход к пределу  $l_v \rightarrow 0$  является, строго говоря, незаконным, что ставит справедливость вывода теоремы Ван-Циттерта—Цернике на основе (11.18) под сомнение. С теоремой (11.16) мы встретимся еще в одном важном случае, а именно в задаче об излучении поля системой точечных независимых источников, к которой мы обратимся в § 12. Смысл приведенной интенсивности в этой задаче будет, разумеется, иным, чем в (11.17).

Результаты, полученные выше для пространственной корреляции спектральных амплитуд, в случае квазимонохроматического поля сохраняют силу и для самих полей. В самом деле,

<sup>1</sup>) Заметим попутно, что отождествление приведенной интенсивности (11.17) со средней интенсивностью поля в отверстии  $\bar{I}_v(\rho)$  ведет к тому же выводу. Такое отождествление возможно при  $k^2 S_v / 4\pi^2 = 1$ , а это означает, что  $S_v = \lambda^2$ , что противоречит принятой в (11.18) гипотезе о дельта-корреляции, по которой  $S_v \sim l_v^3 \ll \lambda^2$ .

согласно (11.6) пространственная корреляционная функция равна

$$\begin{aligned} \Gamma_{\perp}(\tau=0, \rho_1, \rho_2, z) &= \int_0^{\infty} \Gamma_{\perp}(\omega, \rho_1, \rho_2, z) e^{i\omega\tau} d\omega |_{\tau=0} = \\ &= \int_0^{\infty} d\omega \left( \frac{k}{2\pi z} \right)^2 \exp \left( \frac{ik\rho\rho_+}{z} \right) \int_{-\infty}^{\infty} d^2\xi \Gamma_v(\omega, \xi) \exp \left( -\frac{ik\xi\rho_+}{z} \right) \times \\ &\quad \times \int_{-\infty}^{\infty} d^2\eta M(\eta) \exp \left( -\frac{ik\eta\rho}{z} \right). \quad (11.19) \end{aligned}$$

Для квазимонохроматического поля функция  $\Gamma_v(\omega, \xi)$  сосредоточена в узкой полосе частот  $\Delta\omega$  вблизи частоты  $\omega_0$ . Поэтому приближению ( $k_0 = \omega_0/c$ )

$$\begin{aligned} \Gamma_{\perp}(\tau=0, \rho_1, \rho_2, z) &= \\ &= \left( \frac{k_0}{2\pi z} \right)^2 \exp \left( \frac{ik_0\rho\rho_+}{z} \right) \int_{-\infty}^{\infty} d^2\xi \Gamma_v(\tau=0, \xi) \times \\ &\quad \times \exp \left( -\frac{ik_0\xi\rho_+}{z} \right) \int_{-\infty}^{\infty} d^2\eta M(\eta) \exp \left( -\frac{ik_0\eta\rho}{z} \right), \quad (11.20) \end{aligned}$$

где

$$\Gamma_v(\tau=0, \xi) = \int_0^{\infty} \Gamma_v(\omega, \xi) e^{i\omega\tau} d\omega |_{\tau=0} = \Delta\omega_{\text{эфф}} \Gamma_v(\omega_0, \xi)$$

— пространственная функция корреляции граничного поля, а  $\Delta\omega_{\text{эфф}}$  — эффективная ширина полосы частот.

Переход от (11.19) к (11.20) возможен при выполнении неравенства

$$\left| \frac{\Delta k}{z} (\rho\rho_+ - \xi\rho_+ - \eta\rho) \right| = \frac{\Delta k}{2z} |(\rho_1 - \rho')^2 - (\rho_2 - \rho'')^2| \ll 1, \quad (11.21)$$

физический смысл которого заключается в том, что разность хода волн от разных частей отверстия должна быть мала по сравнению с длиной когерентности  $l_k = c\tau_k \sim c/\Delta\omega$ . Простые оценки показывают, что для выполнения неравенства (11.21) достаточно условия квазимонохроматичности ( $\Delta\omega \ll \omega_0$ ) и условия  $\rho_+ \ll a$ , которое требует, чтобы «центр тяжести» точек наблюдения  $\rho_+ = 1/2(\rho_1 + \rho_2)$  не выходил за пределы размеров отверстия  $a$ .

2. Фокусировка случайных волн. Дифракционную картину в фокусе линзы можно получить из предыдущих результатов при помощи простых преобразований, поскольку поле

в фокальной плоскости линзы подобно полю отверстия на бесконечности.

Пусть в отверстии непрозрачного экрана находится тонкая линза с главным фокусным расстоянием  $F$ . Действие линзы можно описать, введя под знак интеграла в (11.3) фазовый множитель  $\exp\{-ik\rho'^2/2F\}$ :

$$u(\mathbf{r}) = \frac{keikz}{2\pi iz} \int_{-\infty}^{\infty} M(\rho') v(\rho') \exp\left(-\frac{ik\rho'^2}{2F}\right) \exp\left(\frac{ik(\mathbf{r}-\rho')^2}{2z}\right) d^2\rho'. \quad (11.22)$$

В главной фокальной плоскости  $z=F$  показатель экспоненты упрощается и поле  $u(\rho, F)$  выражается просто преобразованием Фурье от  $M(\rho') v(\rho')$ :

$$u(\rho, F) = \frac{k \exp\left[ik\left(F + \frac{\rho^2}{2F}\right)\right]}{2\pi i F} \int_{-\infty}^{\infty} M(\rho') v(\rho') \exp\left(-\frac{ik\rho\rho'}{F}\right) d^2\rho'. \quad (11.23)$$

От фразунгоферова приближения (11.11) для поля отверстия (без линзы) выражение (11.23) отличается только заменой  $z$  на  $F$ , так что формулы для функций корреляции повторяют соответствующие выражения из п. 1. Так, при  $a \ll l_v$  (малое отверстие) распределение интенсивности в фокальной плоскости оказывается таким же, как и для пространственно когерентной плоской волны. В противоположном случае  $a \gg l_v$  (большое отверстие) распределение интенсивности может быть получено из (11.13):

$$\bar{I}_u(\rho_+, F) = \frac{k^2 S}{(2\pi F)^2} \int_{-\infty}^{\infty} \Gamma_v(\xi) \exp\left(-\frac{ik\rho_+\xi}{F}\right) d^2\xi \quad (a \gg l_v).$$

Размер фокального пятна в этом случае дается выражением  $\rho_\Phi \sim \lambda F/l_v$ , т. е. оно в  $a/l_v$  раз больше, чем для детерминированного поля плоской волны ( $\rho_\Phi \sim \lambda F/a$ ). Однако интенсивность в центре дифракционной картины теперь равна

$$\bar{I}_u(\rho_+ = 0, F) \approx \frac{k^2 S S_v \bar{I}_v}{(2\pi F)^2},$$

что примерно в  $S/S_v \sim (a/l_v)^2$  раз меньше, чем при когерентном освещении (малое отверстие). Разумеется, полная интенсивность, полученная интегрированием  $\bar{I}_u$  по всей плоскости  $z=F$ , в обоих случаях одинакова и равна полной «входной» интенсивности  $\bar{I}_v S$ .



Как и при дифракции на отверстии, распределение интенсивности для квазимонохроматического поля оказывается практически таким же, как и для строго монохроматической волны. Более того, в оптике даже в случае *белого света*, у которого  $\Delta\omega \sim \omega_0$ , распределение интенсивности в плоскости  $z = F$  лишь незначительно отличается от соответствующей картины для монохроматического поля. Именно поэтому при рассмотрении задачи формирования оптического изображения допустимо пренебрежение эффектами, обусловленными временной некогерентностью поля [21, 36].

3. О роли пространственной когерентности освещения в формировании оптического изображения. Рассмотрим простейшую оптическую систему, содержащую только одну тонкую линзу (рис. 7). Полупрозрачный

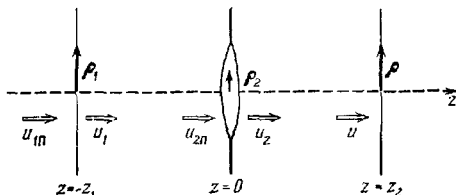


Рис. 7.

объект, например диапозитив, характеризуемый комплексным коэффициентом пропускания  $f(\rho_1)$ , расположен в предметной плоскости  $z = -z_1$ . При освещении этого объекта слева плоской монохроматической волной  $u_{\text{пад}} = u_{\text{ин}}(\rho, z)$  непосредственно за объектом создается поле

$$u_1(\rho_1) = f(\rho_1) v(\rho_1), \quad v(\rho_1) \equiv u_{\text{ин}}(\rho, z)|_{z=-z_1}, \quad (11.24)$$

где  $\rho_1$  — радиус-вектор точки в предметной плоскости.

В плоскости  $z = 0$ , где расположена линза, образуется поле  $u_{2n}(\rho_2)$  ( $\rho_2$  — вектор в плоскости  $z = 0$ ), которое во френелевском приближении (9.17) равно

$$u_{2n}(\rho_2) = \frac{ke^{ikz_1}}{2\pi iz_1} \int_{-\infty}^{\infty} u_1(\rho_1) \exp \left[ \frac{ik(\rho_2 - \rho_1)^2}{2z_1} \right] d^2\rho_1. \quad (11.25)$$

«На выходе» линзы возникает поле

$$u_2(\rho_2) = u_{2n}(\rho_2) M(\rho_2) \exp \left( -\frac{ik\rho_2^2}{2F} \right), \quad (11.26)$$

где множитель  $M(\rho_2)$ , равный нулю вне линзы и единице на линзе, описывает действие ограничивающей линзу диафрагмы, а множитель  $\exp\{-ik\rho_2^2/2F\}$  — действие линзы с фокусным расстоянием  $F$ .

Наконец, пересчет поля (11.26) от линзы к плоскости изображения  $z=z_2$  также может быть произведен во френелевском приближении:

$$u(\rho) = \frac{ke^{ikz_2}}{2\pi iz_2} \int_{-\infty}^{\infty} u_2(\rho_2) \exp\left(\frac{ik(\rho - \rho_2)^2}{2z_2}\right) d^2\rho_2. \quad (11.27)$$

Выражения (11.24)–(11.27) связывают поле  $u(\rho)$  в плоскости изображения с коэффициентом пропускания  $f(\rho_1)$ , характеризующим объект, и с первичным полем  $v(\rho_1)$ , свойства которого определяются источником света. Не рассматривая собственно оптический аспект задачи, уделим основное внимание особенностям формирования изображения в зависимости от степени пространственной когерентности освещения полупрозрачного объекта.

С этой целью примем с самого начала, что плоскости предмета и изображения сопряжены в смысле геометрической оптики, т. е.  $1/z_1 + 1/z_2 = 1/F$ . Подставив последовательно (11.26), (11.25) и (11.24) в (11.27), получаем

$$u(\rho) = \frac{-k^2}{4\pi^2 z_1 z_2} \exp\left[ik\left(z_1 + z_2 + \frac{\rho^2}{2z_2}\right)\right] \times \\ \times \int_{-\infty}^{\infty} \int f(\rho_1) v(\rho_1) M(\rho_2) \exp\left[-ik\rho_2\left(\frac{\rho}{z_2} + \frac{\rho_1}{z_1}\right) + ik\frac{\rho_1^2}{2z_1}\right] d^2\rho_1 d^2\rho_2.$$

Интегрирование по  $\rho_2$  дает

$$\int M(\rho_2) \exp\left[-ik\rho_2\left(\frac{\rho}{z_2} + \frac{\rho_1}{z_1}\right)\right] d^2\rho_2 = S \mathcal{F}_M\left[k\left(\frac{\rho}{z_2} + \frac{\rho_1}{z_1}\right)\right],$$

где  $S$  — площадь отверстия, а через  $\mathcal{F}_M(x)$  обозначена нормированная к единице ( $\mathcal{F}_M(0) = 1$ ) трансформанта Фурье от  $M(\rho)$  (см. (11.14)). В результате

$$u(\rho) = \\ = -\frac{k^2 S \exp[ik(z_1 + z_2 + \rho^2/2z_2)]}{4\pi^2 z_1 z_2} \int_{-\infty}^{\infty} v(\rho_1) f(\rho_1) \mathcal{F}_M\left[k\left(\frac{\rho}{z_2} + \frac{\rho_1}{z_1}\right)\right] \times \\ \times \exp\left(\frac{ik\rho_1^2}{2z_1}\right) d^2\rho_1. \quad (11.28)$$

Выражение (11.28) представляет собой частный случай более общей формулы

$$u(\rho) = \int v(\rho_1) f(\rho_1) \mathcal{P}(\rho, \rho_1) d^2 \rho_1, \quad (11.29)$$

которая принадлежит к соотношениям типа (9.2) и в рассматриваемой задаче связывает поле в предметной плоскости произвольной оптической системы с облучающим полем  $v(\rho_1)$  и функцией пропускания объекта  $f(\rho_1)$ . Функцию  $\mathcal{P}(\rho, \rho_1)$  называют *аппаратной функцией* системы. Характерный масштаб изменения аппаратной функции в плоскости изображения (т. е. по  $\rho$ ) — это радиус дифракционного пятна, отвечающего точечному объекту, а масштаб изменения  $\mathcal{P}$  по аргументу  $\rho_1$  определяет размер области в плоскости предмета, дающей заметный вклад в поле в данной точке плоскости изображения. Иными словами, это предел разрешения оптической системы.

В рассматриваемом случае однолинзовой системы с круглой диафрагмой аппаратная функция равна

$$\mathcal{P}(\rho, \rho_1) = - \frac{k^2 S \exp[ik(z_1 + z_2 + \rho^2/2z_2 + \rho_1^2/2z_1)]}{4\pi^2 z_1 z_2} \frac{2J_1(\kappa a)}{\kappa a}, \quad (11.30)$$

где  $\kappa = k|\rho/z_2 + \rho_1/z_1|$ . Согласно (11.30) точечному объекту, помещенному в точку  $\rho_{10}$ , отвечает дифракционное пятно с центром в точке  $\rho_0 = -\rho_{10} z_2/z_1$ . Отношение  $z_2/z_1$  характеризует увеличение данной системы, а знак минус отвечает перевернутому изображению. Так как функция Бесселя  $J_1(\kappa a)$  первый раз обращается в нуль при  $\kappa a = 3,83$ , характерным масштабом аппаратной функции по  $\rho$  может служить радиус первого темного кольца Эйри в дифракционном изображении точечного объекта, равный  $\frac{3,83}{\kappa a} z_2 \approx 0,61 \lambda \frac{z_2}{a}$ , тогда как масштаб изменения по  $\rho_1$  (с точностью до множителя порядка единицы — это релеевский предел разрешения) равен  $\delta \equiv 0,61 \lambda \frac{z_1}{a}$ .

При анализе свойств изображения мы будем исходить из общей формулы (11.29), а частный вид этой формулы (11.28) используем только для иллюстрации общих выводов. Согласно (11.29) при освещении объектов частично когерентным светом средняя интенсивность поля в плоскости изображения равна

$$\bar{I}(\rho) = \int \Gamma_v(\rho', \rho'') f(\rho') f^*(\rho'') \mathcal{P}(\rho, \rho') \mathcal{P}^*(\rho, \rho'') d^2 \rho' d^2 \rho'', \quad (11.31)$$

где

$$\Gamma_v(\rho', \rho'') = \langle v(\rho') v^*(\rho'') \rangle$$

— пространственная функция когерентности первичного поля.

Для заданного объекта (функция  $f$  фиксирована) характер распределения интенсивности в предметной плоскости зависит

от соотношения между радиусом когерентности первичного поля  $l_v$  (это характерный масштаб изменения функции когерентности  $\Gamma_v$ ), размером области разрешения  $\delta$  (характерный масштаб изменения аппаратной функции по  $\rho_1$ ) и размером наиболее мелких деталей объекта  $\rho_{\min}$ . Соотношение между  $\rho_{\min}$  и  $\delta$  определяется характеристиками соответственно объекта и оптического прибора. Если считать эти величины заданными, то воздействовать на изображение можно, только меняя  $l_v$ , т. е. меняя характер облучения.

Рассмотрим предельные случаи  $l_v \gg \rho_{\min}$ ,  $\delta$  и  $l_v \ll \rho_{\min}$ ,  $\delta$ , которые принято называть соответственно когерентным и некогерентным освещением предмета. При когерентном освещении ( $l_v \gg \rho_{\min}$ ,  $\delta$ ) можно считать, что в области, существенной для интегрирования в (11.31) (в этой области обе функции  $\mathcal{P}(\rho, \rho')$  и  $\mathcal{P}^*(\rho, \rho'')$  отличны от нуля, так что  $|\rho' - \rho''| \leq \delta$ ), функция когерентности  $\Gamma_v(\rho' - \rho'')$  практически постоянна (поскольку  $|\rho' - \rho''| \leq \delta \ll l_v$ ). Вынося ее за знак интеграла со значением  $\bar{I}_v$ , где  $\bar{I}_v = \Gamma_v(0)$  — интенсивность облучающего поля, получаем

$$\bar{I}(\rho)_{\text{ког}} = \bar{I}_v \left| \int_{-\infty}^{\infty} f(\rho_1) \mathcal{P}(\rho, \rho_1) d^2 \rho_1 \right|^2. \quad (11.32)$$

Такую же формулу можно получить, разумеется, и непосредственно из (11.31), если считать первичное поле  $v(\rho_1)$  когерентным на всей плоскости предмета. Из приведенного вывода видно, что требование когерентности на всем предмете является излишним: для справедливости (11.32) достаточно, чтобы первичное поле  $v(\rho)$  было когерентным лишь в пределах области разрешения  $\delta$ .

В другом предельном случае некогерентного освещения ( $l_v \ll \ll \delta, \rho_{\min}$ ) существенная для интегрирования область в (11.32) ограничена площадью когерентности с линейными размерами  $|\rho' - \rho''| \leq l_v$ . В пределах этой площади функции  $f^*(\rho'')$  и  $\mathcal{P}^*(\rho, \rho'')$  можно приближенно заменить соответственно на  $f^*(\rho')$  и  $\mathcal{P}^*(\rho, \rho')$ , и в результате получим

$$\bar{I}(\rho)_{\text{неког}} = \int_{-\infty}^{\infty} |f(\rho') \mathcal{P}(\rho, \rho')|^2 d^2 \rho' \int_{-\infty}^{\infty} \Gamma_v(\rho' - \rho'') d^2 \rho''.$$

Внутренний интеграл равен  $\bar{I}_v S_v$ , где, подобно (11.15), введена площадь когерентности  $S_v$ , равная интегралу от коэффициента корреляции первичного поля  $K_v = \Gamma_v / \bar{I}_v$ .

Таким образом, распределение средней интенсивности в плоскости изображения при некогерентном освещении дается выражением

$$\bar{I}(\rho)_{\text{неког}} = \bar{I}_v S_v \int |f(\rho_1)|^2 |\mathcal{P}(\rho, \rho_1)|^2 d^2 \rho_1. \quad (11.33)$$

Формально это выражение можно получить из (11.31), считая, как и в (11.18), первичное поле дельта-коррелированным.

Формулы (11.32) и (11.33) существенно отличаются друг от друга. Можно сказать, что при когерентном (в указанном выше смысле, т. е. при  $l_v \gg \delta$ ,  $\rho_{\min}$ ) освещении происходит преобразование самого комплексного коэффициента пропускания  $f(\rho_1)$  (с передаточной функцией  $\sqrt{\bar{I}_v} \mathcal{P}(\rho, \rho_1)$ ), а при некогерентном освещении ( $l_v \ll \delta$ ,  $\rho_{\min}$ ) преобразуется квадрат модуля  $|f(\rho_1)|^2$  (с передаточной функцией  $\bar{I}_v S_v |\mathcal{P}(\rho, \rho_1)|^2$ ). Отсюда следует, в частности, что при некогерентном освещении изображение не зависит от фазы комплексной функции пропускания  $f$ , тогда как когерентное освещение может выявить резкие вариации фазы  $f$  даже при  $|f| = \text{const}$ . Например, если на некоторой линии в плоскости предмета фаза скачком меняется на  $\pi$ , а  $|f| = \text{const}$ , то, как следует из (11.32), в плоскости изображения на соответствующей (сопряженной) линии интенсивность  $\bar{I}(\rho)_{\text{ког}}$  обратится в нуль и мы увидим там темную полосу. Правда, при *плавном* (в масштабе  $\delta$ ) изменении модуля и аргумента коэффициента прозрачности, т. е. при  $\rho_{\min} \gg \delta$ , оба вида освещения дают практически одинаковое изображение, так как в обоих случаях распределение интенсивности в плоскости изображения  $\bar{I}(\rho)$  пропорционально  $|f(\rho_1)|^2$ .

Укажем еще на два различия между когерентным и некогерентным освещением. Во-первых, при одинаковой интенсивности освещения  $I_v$  интенсивность в плоскости изображения при когерентном свете будет больше, чем при некогерентном, поскольку ширина углового спектра у некогерентного облучения больше, чем у когерентного. Во-вторых, в предельных случаях освещения возникают различия и в величине разрешающей способности: при некогерентном освещении размер наименьших различимых по Релею деталей несколько меньше (примерно на 30%), чем при когерентном. Это различие связано просто с тем, что входящий в формулу (11.33) квадрат модуля  $|\mathcal{P}(\rho, \rho_1)|^2$  меняется в функции от координаты  $\rho_1$  круче, чем сама аппаратная функция  $\mathcal{P}(\rho, \rho_1)$  в формуле (11.32). В результате при разрешении двух точечных объектов с центрами в точках  $\rho'_1$  и  $\rho''_1$  провал между горбами в распределении интенсивности в плоскости изображения получается более глубоким, если световые колебания в точках  $\rho'_1$  и  $\rho''_1$  некогерентны. Провал можно сделать еще глубже, если добиться *отрицательной* корреляции освещающего поля в точках  $\rho'_1$  и  $\rho''_1$  (отрицательные значения может принимать, например, функция когерентности вида  $2J_1(x)/x$ , отвечающая равномерно светящемуся диску или же пучку света, пропущенному через малое круглое отверстие). Этот красивый по своей идее способ увеличения разрешающей способности дает, однако, весьма незначительный выигрыш (всего лишь на несколько

процентов) по сравнению с некогерентным облучением. Более подробно этот вопрос рассмотрен в книге [21].

Возникающие в оптике статистические задачи, конечно, не исчерпываются приведенными примерами. Среди других проблем упомянем пространственную фильтрацию, статистические вопросы, связанные с несовершенством оптических систем и с зернистостью фотоматериалов, использование оптических систем для корреляционного анализа и т. д. С этими и некоторыми другими приложениями можно познакомиться по работам [13—22]. Многие идеи, развитые первоначально в оптике, нашли применение и в радиотехнике [37].

## § 12. Возбуждение полей случайными источниками

Флуктуации поля, возбуждаемого случайными источниками, представляют интерес для многих разделов физики. Здесь мы остановимся на излучении радиоантенн со случайными вариациями токов в раскрыве и рассмотрим особенности излучения большого числа независимых источников.

1. Статистика поля, излученного большой антенной. Причиной флуктуаций поля, излученного антенной, могут быть случайные отклонения амплитуд и фаз токов в апертуре от заданных значений, обусловленные различными факторами. Главными из них являются неровности антенного зеркала, разброс параметров излучающих вибраторов, случайные отклонения в системах возбуждения и т. д. Определенную роль играют также деформации антенн при изменении температуры или из-за ветровых нагрузок. Рассмотрим одну из моделей антенны с флуктуациями токов.

Если размеры плоской антенны велики по сравнению с длиной волны, а токи в раскрыве монохроматические ( $\sim e^{-i\omega t}$ ) и имеют только одну компоненту, скажем  $j_x$ , то в направлениях, не сильно отклоняющихся от нормали к плоскости раскрыва  $z=0$ , можно считать, что электрическое поле имеет только компоненту  $E_x$ , удовлетворяющую уравнению

$$\Delta E_x + k^2 E_x = -\frac{4\pi i k}{c} j_x.$$

Таким образом, с учетом иных обозначений ( $u = E_x$ ,  $q = -\frac{4\pi i k}{c} j_x$ ) задача сводится к уже рассмотренной скалярной задаче (9.7), имеющей решение (9.8).

В случае антенн с плоским раскрывом токи сосредоточены в плоскости  $z=0$ , т. е. объемный интеграл (9.8) превращается в поверхностный. Сохраняя для плотности поверхностных токов

обозначение  $q$ , получаем

$$u(\mathbf{r}) = \frac{1}{4\pi} \int_S q(\mathbf{p}') \frac{e^{ikR}}{R} d^2\rho', \quad R = |\mathbf{r} - \mathbf{r}'|, \quad (12.1)$$

где для краткости опущен аргумент  $\omega$ , а интегрирование распространяется на площадь раскрыва  $S$ .

Обычно представляет интерес поле в дальней (фраунгоферовой) зоне  $z \gg ka^2$ , где  $a$  — поперечник раскрыва. Можно поэтому воспользоваться приближением фраунгоферовой дифракции, т. е. формулой (9.10), которая в данном случае (поверхностные источники, излучение сосредоточено в узком конусе около оси  $z$ ) принимает вид

$$u(\rho, z) = \frac{1}{4\pi z} \exp \left[ ik \left( z + \frac{\rho^2}{2z} \right) \right] \int_S q(\rho') \exp \left( -\frac{ik\rho\rho'}{z} \right) d^2\rho'. \quad (12.2)$$

Здесь  $\rho = (x, y)$  — координаты точки наблюдения на удаленной плоскости  $z = \text{const}$ . Ограничимся нахождением среднего значения к пространственной функции корреляции поля, излученного большой антенной, имея в виду лишь основные эффекты, к которым приводят флуктуации токов в антеннах. Систематическому изложению вопросов статистической теории антенн посвящена книга [38].

Как и в § 11, удобно ввести функцию  $M(\rho)$ , которая равна единице на раскрыве  $S$  и обращается в нуль вне  $S$  (см. (11.2)):

$$q(\rho) = M(\rho) [\bar{q}(\rho) + \tilde{q}(\rho)],$$

что позволяет раздвинуть пределы интегрирования в (12.2) до бесконечности:

$$\begin{aligned} u(\rho, z) &= \\ &= \frac{1}{4\pi z} \exp \left[ ik \left( z + \frac{\rho^2}{2z} \right) \right] \int_{-\infty}^{\infty} M(\rho') [\bar{q}(\rho') + \tilde{q}(\rho')] \exp \left( -\frac{ik\rho\rho'}{z} \right) d^2\rho'. \end{aligned} \quad (12.3)$$

Среднее по ансамблю реализаций токов значение поля дается интегралом

$$\bar{u}(\rho, z) = \frac{1}{4\pi z} \exp \left[ ik \left( z + \frac{\rho^2}{2z} \right) \right] \int_{-\infty}^{\infty} \bar{q}(\rho') M(\rho') \exp \left( -\frac{ik\rho\rho'}{z} \right) d^2\rho',$$

который можно представить в виде

$$\bar{u}(\rho, z) = \frac{1}{4\pi z} \exp \left[ ik \left( z + \frac{\rho^2}{2z} \right) \right] \bar{q}_m S_A \mathcal{F}_A \left( \frac{k\rho}{z} \right), \quad (12.4)$$

где  $\bar{q}_m$  — максимальное значение среднего тока в раскрыве, а  $\mathcal{F}_A(\mathbf{x})$  — нормированная к единице ( $\mathcal{F}_A(0)=1$ ) диаграмма направленности антенны:

$$\mathcal{F}_A(\mathbf{x}) = \frac{1}{S_A} \int_{-\infty}^{\infty} M(\rho') \frac{\bar{q}(\rho')}{\bar{q}_m} \exp(i\mathbf{x}\rho') d^2\rho', \quad (12.5)$$

причем направление на точку наблюдения  $(x, y)$  в плоскости  $z = \text{const}$  задается вектором  $\mathbf{x} = k\rho/z = (kx/z, ky/z)$ . Наконец,

$$S_A = \frac{1}{q_m} \int M(\rho') \bar{q}(\rho') d^2\rho'$$

— нормировочный множитель, значение которого близко к площади раскрыва ( $S_A = S$  при  $\bar{q} = \text{const}$ ).

Функция корреляции поля в плоскости  $z = \text{const}$  (поперечная функция корреляции) выражается двукратным интегралом

$$\begin{aligned} \psi_{\perp}(\rho_1, \rho_2, z) &= \langle \tilde{u}(\rho_1, z) \tilde{u}^*(\rho_2, z) \rangle = \\ &= \frac{\exp[ik(\rho_1^2 - \rho_2^2)/2z]}{(4\pi z)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} M(\rho') M(\rho'') \psi_q(\rho', \rho'') \times \\ &\times \exp\left[\frac{ik(\rho_1\rho' - \rho_2\rho'')}{z}\right] d^2\rho' d^2\rho'', \quad (12.6) \end{aligned}$$

где  $\psi_q(\rho', \rho'') = \langle \bar{q}(\rho') \bar{q}^*(\rho'') \rangle$  — функция корреляции тока. Вообще говоря, флуктуации тока в раскрыве статистически неоднородны, что связано с краевыми эффектами: вследствие взаимного влияния элементов излучающей системы статистика токов на краях раскрыва может отличаться от статистики в середине антенны. Однако в случае больших антенн, размеры которых  $a$  велики по сравнению с радиусом корреляции токов  $l_q$ , краевым эффектом можно пренебречь и считать, что на всем раскрыве

$$\psi_q(\rho', \rho'') \approx \psi_q(\rho' - \rho''), \quad l_q \ll a.$$

Принимая модель мелкомасштабных ( $l_q \ll a$ ) флуктуаций тока в раскрыве, мы можем воспользоваться изложенными в § 11 приемами вычисления интегралов типа (11.5). Если обозначить  $\rho = \rho_1 - \rho_2$ ,  $\rho_+ = (\rho_1 + \rho_2)/2$ , то (12.6) примет вид

$$\begin{aligned} \psi_{\perp}(\rho, \rho_+, z) &= \\ &= \frac{\exp(ik\rho\rho_+/z)}{(4\pi z)^2} \sigma_q^2 \int_{-\infty}^{\infty} M(\eta) \exp\left(\frac{ik\rho\eta}{z}\right) d^2\eta \int_{-\infty}^{\infty} K_q(\xi) \exp\left(\frac{ik\rho_+\xi}{z}\right) d^2\xi, \quad (12.7) \end{aligned}$$

где  $\sigma_q^2 = \langle |\bar{q}|^2 \rangle$  — дисперсия, а  $K_q$  — коэффициент корреляции флуктуаций тока.



Первый интеграл в (12.7) равен  $S\mathcal{F}_M\left(\frac{k\rho}{z}\right)$ , где  $S$ —площадь раскрыва, а  $\mathcal{F}_M$  дается выражением (11.14). Эта величина несколько отличается от диаграммы антенны (12.5), но при  $\bar{q} = \text{const}$   $\mathcal{F}_M = \mathcal{F}_A$ . Второй интеграл (по  $\xi$ ) тоже можно записать в виде произведения двух множителей:

$$\int_{-\infty}^{\infty} K_q(\xi) \exp\left(\frac{ik\rho_+\xi}{z}\right) d^2\xi = S_q \mathcal{F}_q\left(\frac{k\rho_+}{z}\right), \quad (12.8)$$

где

$$S_q = \int_{-\infty}^{\infty} K_q(\xi) d^2\xi$$

—эффективная площадь когерентности токов, а  $\mathcal{F}_q(\kappa)$ —нормированная к единице ( $\mathcal{F}_q(0)=1$ ) диаграмма направленности флуктуационных токов.

С введением этих обозначений корреляционная функция  $\psi_{\perp}$  принимает вид

$$\psi_{\perp}(\rho, \rho_+, z) = \exp\left(ik\frac{\rho\rho_+}{z}\right) \frac{\sigma_u^2 S S_q}{(4\pi z)^2} \mathcal{F}_M\left(\frac{k\rho}{z}\right) \mathcal{F}_q\left(\frac{k\rho_+}{z}\right). \quad (12.9)$$

Эту формулу можно рассматривать как обобщение теоремы Ван-Циттерта—Цернике на случай конечного радиуса корреляции источников. При дельта-коррелированных флуктуациях  $\mathcal{F}_q=1$ , и тогда (12.9) отличается от (11.16) только обозначениями. Отсюда следует, что пространственное распределение дисперсии поля определяется сравнительно широкой (с угловым раствором  $\sim \lambda/l_q$ ) диаграммой  $\mathcal{F}_q$ :

$$\sigma_u^2(\rho_+, z) = \psi_{\perp}(0, \rho_+, z) = \frac{\sigma_u^2 S S_q}{(4\pi z)^2} \mathcal{F}_q\left(\frac{k\rho_+}{z}\right), \quad (12.10)$$

тогда как пространственный коэффициент корреляции поля зависит от диаграммы  $\mathcal{F}_M$ :

$$K_{\perp}(\rho, z) = \frac{\psi_{\perp}(\rho, \rho_+, z)}{\sigma_u^2(\rho_+, z)} = \exp\left(ik\frac{\rho\rho_+}{z}\right) \mathcal{F}_M\left(\frac{k\rho}{z}\right).$$

Последняя имеет угловую ширину  $\sim \lambda/a$ , меньшую, чем ширина  $\lambda/l_q$  флуктуационной диаграммы  $\mathcal{F}_q$ , поскольку  $a \gg l_q$ . Поперечный радиус корреляции  $l_{\perp}$  оказывается таким же, как в (11.12), т. е.  $l_{\perp} \sim \lambda z/a$ .

Пространственное распределение средней интенсивности  $\bar{I}_u$  можно получить при помощи (12.4) и (12.10):

$$\bar{I}_u(\rho_+, z) = [\bar{u}(\rho_+, z)]^2 \div \sigma_u^2(\rho_+, z) = \frac{1}{(4\pi z)^2} \left[ \bar{q}_M^2 S_A^2 \left| \mathcal{F}_A\left(\frac{k\rho_+}{z}\right) \right|^2 + \sigma_q^2 S_q S \mathcal{F}_q\left(\frac{k\rho_+}{z}\right) \right]. \quad (12.11)$$

Здесь первое слагаемое описывает вклад среднего поля, а второе — вклад флуктуационного поля. На рис. 8 показана зависимость  $\bar{I}_u$  от  $\rho_+$  для различных соотношений между этими слагаемыми в предположении, что полная мощность, подводимая к антенне, постоянна, т. е.

$$|\bar{q}_m|^2 + \sigma_q^2 = \text{const} \equiv Q^2. \quad (12.12)$$

В отсутствие флуктуаций, когда  $\sigma_q = 0$  и  $|\bar{q}_m| = Q$ , мы имеем невозмущенную интенсивность

$$I_0(\rho_+, z) = \frac{S_A^2 Q^2}{(4\pi z)^2} \left| \mathcal{F}_A \left( \frac{k\rho_+}{z} \right) \right|^2, \quad (12.13)$$

пропорциональную квадрату модуля диаграммы направленности  $\mathcal{F}_A$ . Ширина этой диаграммы, скажем расстояние между ее нулями, составляет  $\lambda z/a$  (кривая 1 на рис. 8). При малых

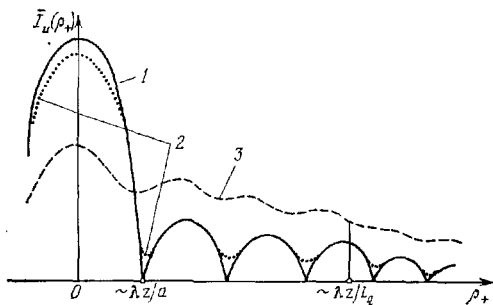


Рис. 8.

возмущениях появляется флуктуационная компонента, пропорциональная угловому спектру флуктуаций  $\mathcal{F}_q(k\rho_+/z)$  и имеющая ширину  $\Delta\rho_+ \sim \lambda z/l_q \gg \lambda z/a$ . Эта составляющая приводит к «замыванию» нулей (или минимумов) регулярной диаграммы (кривая 2 на рис. 8) и к росту бокового излучения, т. е. излучения вне центрального максимума. Вместе с тем несколько уменьшается вклад регулярной составляющей поля, так как при неизменной мощности, подводимой к антенне, величина  $|\bar{q}_m|^2$  уменьшается по сравнению с  $Q^2$ :  $|\bar{q}_m|^2 = Q^2 - \sigma_q^2$ .

Распределение интенсивности в плоскости  $z = \text{const}$  в случае примерного равенства регулярной и флуктуационной компонент в центре диаграммы изображает кривая 3 на рис. 8. Согласно

(12.11) равенство  $|\bar{u}|^2 = \sigma_u^2$  достигается при  $\sigma_q^2 = Q^2 \left(1 - \frac{SS_q}{S_A^2}\right)$ ;

при этом квадрат среднего значения тока  $|\bar{q}_m|^2 = Q^2 - \sigma_q^2$  мал по сравнению с  $Q^2$ , т. е.  $|\bar{q}_m|^2 \approx \frac{SS_q}{S_A^2} Q^2 \ll Q^2$ . При столь больших возмущениях говорят о *разрушении* диаграммы направленности антенны.

С дальнейшим ростом флуктуаций регулярное слагаемое в (12.11) становится пренебрежимо малым и распределение средней интенсивности уже совсем не похоже на невозмущенное распределение (12.13):

$$\bar{I}_u(\rho_+, z) \approx \sigma_u^2(\rho_+, z) = \frac{SS_q}{(4\pi z)^2} \sigma_q^2 \mathcal{F}_q\left(\frac{k\rho_+}{z}\right).$$

Ширина диаграммы излучения здесь порядка  $\lambda z/l_q$ , а максимум интенсивности в  $S_q/S$  раз меньше, чем в отсутствие флуктуаций.

Уменьшение интенсивности в направлении максимума излучения удобно характеризовать величиной *коэффициента направленного действия* (к. н. д.), который показывает, во сколько раз интенсивность поля в центре главного лепестка диаграммы больше, чем у ненаправленной антенны, излучающей ту же мощность. Вычислим средний к. н. д. антенны  $\langle G \rangle$  при сделанных выше допущениях. Очевидно, если  $G_0$  — к. н. д. невозмущенной антенны, то

$$\langle G \rangle = G_0 \frac{\bar{I}_{\max}}{I_{0\max}}.$$

Так как

$$I_{0\max} = \frac{S_A^2 Q^2}{(4\pi z)^2}, \quad \bar{I}_{\max} = \frac{1}{(4\pi z)^2} [|\bar{q}_m|^2 S_A^2 + SS_q \sigma_q^2],$$

для величины  $\langle G \rangle / G_0$  с учетом (12.12) получаем

$$\frac{\langle G \rangle}{G_0} = \frac{|\bar{q}_m|^2 + (S_q S / S_A^2) \sigma_q^2}{Q^2} = 1 - \left(1 - \frac{SS_q}{S_A^2}\right) \frac{\sigma_q^2}{Q^2}. \quad (12.14)$$

При малых возмущениях токов в раскрыве ( $\sigma_q \ll Q$ ) средний к. н. д. антенны мало отличается от  $G_0$ :  $\langle G \rangle \approx G_0 (1 - \sigma_q^2 / Q^2)$  (напомним, что  $S_q \ll S$ ). Уменьшение среднего к. н. д. вдвое наступает уже при заметных флуктуациях, когда  $\sigma_q^2 \sim 1/2 Q^2$ . При сильных же флуктуациях, когда  $\sigma_q \rightarrow Q$ , а регулярная составляющая тока  $|\bar{q}_m| = \sqrt{Q^2 - \sigma_q^2}$  стремится к нулю, получаем

$$\frac{\langle G \rangle}{G_0} \approx \frac{S_q S}{S_A^2} \sim \frac{S_q}{S_A}. \quad (12.15)$$

Таким образом, при сильных флуктуациях к. н. д. антенны тем меньше, чем больше неоднородностей  $N = S/S_q$  содержит раскрыв:  $\langle G \rangle = G_0/N$ . Учитывая, что  $G_0 \sim S/\lambda^2$ , соотношение (12.15) можно записать иначе:

$$\langle G \rangle \sim \frac{S_q}{\lambda^2}.$$

Отсюда видно, что если радиус корреляции  $l_q$  сравним с длиной волны  $\lambda$ , то  $S_q \approx l_q^2 \sim \lambda^2$  и  $\langle G \rangle \sim 1$ . Иными словами, при сильных ( $\sigma_q \gg |\bar{q}_m|$ ) и мелких ( $l_q \sim \lambda$ ) флуктуациях излучение антенны полностью теряет направленность.

В статистической теории антенны возможны и многие другие задачи. Например, может представлять интерес направление центрального лепестка диаграммы направленности, уровень бокового излучения, статистика амплитуды и фазы и т. д. Эти вопросы достаточно полно освещены в монографии [38]. Отметим, что при исследовании статистики излучения антенны основная трудность заключается, по-видимому, не в вычислении тех или иных интегралов, входящих в выражения для моментов поля, а в физически обоснованном задании статистики флуктуирующих по раскрыву антенны токов. Ясно, что предсказать эту статистику только из теоретических соображений невозможно. В то же время прямые измерения статистических характеристик сопряжены со значительными трудностями, в частности, из-за того, что при внесении в антенну измерительных зондов распределение токов в ней меняется. Косвенное же нахождение флуктуаций путем измерения полей, излученных антенной, тоже не дает исчерпывающего решения проблемы: такие измерения подвержены влиянию дополнительных посторонних факторов (расположенные рядом предметы и постройки, неровности местности и т. д.), не говоря уже о том, что измерения поля сложны сами по себе как в ближней, так и в дальней зонах.

2. Статистика поля, излученного системой независимых источников. Для простоты будем считать волновое поле скалярным, а источники — точечными. Если  $j$ -й точечный источник движется по траектории  $\mathbf{r} = \mathbf{r}_j(t)$ , а его интенсивность меняется по закону (вообще говоря, случайному)  $Q_j(t)$ , то для системы  $N$  точечных источников правая часть волнового уравнения (9.4) запишется в виде

$$q(t, \mathbf{r}) = \sum_{j=1}^N Q_j(t) \delta[\mathbf{r} - \mathbf{r}_j(t)].$$

В результате поле излучения (9.5) будет выражаться суммой

$$u(t, \mathbf{r}) = \frac{1}{4\pi} \sum_{j=1}^N Q_j \left( t - \frac{|\mathbf{r} - \mathbf{r}_j(t)|}{c} \right) \frac{1}{|\mathbf{r} - \mathbf{r}_j(t)|}, \quad (12.16)$$

в которой  $|r - r_j(t)|$  — расстояние от точки наблюдения до  $j$ -го источника.

Не интересуясь средним значением поля, которое в подавляющем большинстве задач равно нулю, вычислим пространственно-временную функцию когерентности, сначала — для системы неподвижных источников ( $r_j = \text{const}$ ). В силу (12.16)

$$\Gamma(t', r'; t'', r'') = \langle u(t', r') u^*(t'', r'') \rangle =$$

$$= \frac{1}{(4\pi)^2} \sum_{j=1}^N \sum_{l=1}^N \left\langle \frac{Q_j \left( t' - \frac{1}{c} |r' - r_j| \right) Q_l^* \left( t'' - \frac{1}{c} |r'' - r_l| \right)}{|r' - r_j| |r'' - r_l|} \right\rangle, \quad (12.17)$$

где угловые скобки означают усреднение по ансамблю положений источников, а черта сверху — усреднение по ансамблю реализаций случайных процессов  $Q_j(t)$ .

Выражение (12.17) принимает особенно простой вид для системы независимых и одинаковых источников. Для такой системы при  $j \neq l$

$$\overline{Q_j(t') Q_l^*(t'')} = 0,$$

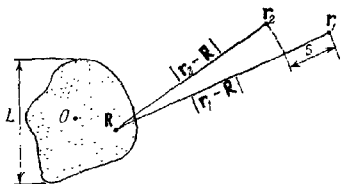


Рис. 9.

если одинаковую для всех процессов  $Q_j(t)$  корреляционную функцию обозначить через  $\psi_Q(\tau)$  (предполагается стационарность этих процессов), то двойная сумма (12.17) превратится в сумму  $N$  одинаковых слагаемых:

$$\Gamma_a(\tau; r_1, r_2) = \frac{N}{(4\pi)^2} \left\langle \frac{\psi_Q(\tau - s/c)}{|r_1 - R| |r_2 - R|} \right\rangle,$$

где  $\tau = t_1 - t_2$ , а  $s = |r_1 - R| - |r_2 - R|$  — разности хода от источника, находящегося в точке  $R$  до точек наблюдения  $r_1$  и  $r_2$  (рис. 9).

Усреднение по положениям источников можно представить в виде интеграла по объему с весовой функцией  $w(R)$ , которая представляет собой плотность вероятностей для радиуса-вектора  $R$ :

$$\Gamma_a(\tau; r_1, r_2) = \frac{N}{(4\pi)^2} \int \frac{\psi_Q(\tau - s/c)}{|r_1 - R| |r_2 - R|} w(R) d^3R,$$

или, если ввести среднюю концентрацию источников  $\bar{n}(R) = Nw(R)$ ,

$$\Gamma_a(\tau; r_1, r_2) = \frac{1}{(4\pi)^2} \int \frac{\psi_Q(\tau - s/c)}{|r_1 - R| |r_2 - R|} \bar{n}(R) d^3R. \quad (12.18)$$

В частном случае, когда источники распределены по объему  $V$  равномерно,  $\bar{n} = N/V = \text{const}$ .

Временная функция когерентности может быть получена из (12.18), если положить  $\mathbf{r}_1 = \mathbf{r}_2 = \mathbf{r}$ . Тогда  $s = 0$  и

$$\Gamma_u(\tau; \mathbf{r}, \mathbf{r}) = \frac{\psi_Q(\tau)}{(4\pi)^2} \int \frac{\bar{n}(\mathbf{R})}{|\mathbf{r} - \mathbf{R}|^2} d^3R.$$

Отсюда видно, что коэффициенты временной корреляции поля и источников совпадают:  $K_u(\tau) = K_Q(\tau)$ , т. е. временная когерентность излученного поля оказывается такой же, как у источников. Это естественно, так как при неподвижных источниках изменения поля обусловлены только изменениями функций  $Q_j(t)$ .

Обращаясь к пространственной корреляции излученного поля, рассмотрим некоторые частные случаи.

Если размеры «облака» источников малы по сравнению со средней длиной волны излучения  $\lambda_0$ ,  $L \sim V^{1/3} \ll \lambda_0$ , то можно считать задержку  $s/c$  в формуле (12.18) постоянной величиной, равной  $s_0/c = (|\mathbf{r}_1 - \mathbf{R}_0| - |\mathbf{r}_2 - \mathbf{R}_0|)/c = (r_1 - r_2)/c$ , где  $\mathbf{R}_0 = 0$  — центр объема  $V$ . Для точек наблюдения  $\mathbf{r}_1$  и  $\mathbf{r}_2$ , лежащих вне объема  $V$ , тогда находим

$$\Gamma_u(\tau; \mathbf{r}_1, \mathbf{r}_2) \approx \frac{N\psi_Q(\tau - s_0/c)}{(4\pi)^2 r_1 r_2},$$

где учтено, что  $\int \bar{n}(\mathbf{R}) d^3R = N$ . При  $\tau = 0$  отсюда получается пространственная функция когерентности излучателя малых размеров:

$$\Gamma_u(0; \mathbf{r}_1, \mathbf{r}_2) = \frac{N\psi_Q(-s_0/c)}{(4\pi)^2 r_1 r_2}, \quad L \ll \lambda_0. \quad (12.19)$$

Согласно (12.19) поле пространственно когерентно внутри шарового слоя толщины  $c\tau_k$ , где  $\tau_k$  — время корреляции источников. Таким образом, пространственный радиус корреляции  $l_k$  совпадает в данном случае с длиной когерентного пучка  $c\tau_k$ .

Иначе обстоит дело для больших (по сравнению с  $\lambda_0$ ) объемов  $V$ . Соответствующие закономерности проще всего установить при помощи величины  $\Gamma_u(\omega; \mathbf{r}_1, \mathbf{r}_2)$ , являющейся преобразованием Фурье по времени от  $\Gamma_u(\tau; \mathbf{r}_1, \mathbf{r}_2)$ <sup>1)</sup>. Для функции когерентности (12.18) имеем [44]

$$\begin{aligned} \Gamma_u(\omega; \mathbf{r}_1, \mathbf{r}_2) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \Gamma_u(\tau, \mathbf{r}_1, \mathbf{r}_2) e^{i\omega\tau} d\tau = \\ &= \frac{1}{4\pi^2} g_Q(\omega) \int \frac{\bar{n}(\mathbf{R}) e^{i\mathbf{k}\mathbf{s}}}{|\mathbf{r}_1 - \mathbf{R}| |\mathbf{r}_2 - \mathbf{R}|} d^3R. \end{aligned} \quad (12.20)$$

<sup>1)</sup> В оптике эту величину называют *взаимной спектральной интенсивностью*, так как она относится к *взаимной* функции когерентности  $\Gamma_u(\tau; \mathbf{r}_1, \mathbf{r}_2)$ , характеризующей статистическую связь поля в двух точках,

Здесь  $k = \omega/c$ , а

$$g_Q(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi_Q(\tau) e^{i\omega\tau} d\tau$$

— спектральная плотность процесса  $Q(t)$ .

Непосредственно из вида формулы (12.20) можно заключить, что если размеры «облака» источников  $L$  велики по сравнению с длиной волны  $\lambda$ , а точки  $r_1$  и  $r_2$  находятся внутри облака, то смещение любой из точек  $r_1$  или  $r_2$  на  $\lambda$  приведет к существенному изменению величины  $\Gamma_u$ , так как под интегралом имеется быстро осциллирующая функция  $e^{iks}$ . Отсюда следует, что внутри большого ( $L \gg \lambda$ ) облака излучателей пространственный радиус корреляции сравним с длиной волны<sup>1)</sup>:

$$l_u \sim \lambda.$$

По мере удаления точек наблюдения от облака должно, очевидно, происходить некоторое упорядочение интерференционной картины — хотя бы из-за того, что волины будут приходить теперь только из ограниченного телесного угла. Это упорядочение проявляется в увеличении пространственного радиуса корреляции вдали от системы излучателей, подобно вытекающему из теоремы Ван-Циттерта — Цернике увеличению  $l_u$  для дифракционного поля отверстия (§ 11). Теорема Ван-Циттерта — Цернике может быть сформулирована и для системы независимых излучателей, но вывод оказывается несколько сложнее, чем в задаче о дифракции случайного поля.

Приведем только окончательный результат. Пусть точки  $r_1$  и  $r_2$  лежат в плоскости  $z = \text{const}$ , при этом  $r_{1,2} = (\rho_{1,2}, z)$ . У вектора  $R$  тоже выделим продольную ( $z'$ ) и поперечную ( $\rho'$ ) координаты:  $R = (\rho', z')$ ,  $d^3R = d^2\rho' dz'$ . Если расстояние от облака велико по сравнению с его поперечником ( $z \gg L$ ), а точки наблюдения расположены не слишком далеко от оси  $z$  ( $\rho_{1,2} \ll z$ ), то формула (12.20) принимает вид

$$\Gamma_u(\omega; \rho, \rho_+, z) = \frac{g_Q(\omega) \exp\left(\frac{ik\rho\rho_+}{z}\right)}{(4\pi z)^2} \int \bar{n}(\rho', z') \exp\left(-\frac{ik\rho\rho'}{z}\right) d^2\rho' dz', \quad (12.21)$$

<sup>1)</sup> Простая иллюстрация: от брошенного в воду камня расходятся волны правильной круговой формы. Если же бросить пригоршню камней, то интерференция круговых волн создает в месте падения камней беспорядочную картину гребней и впадин, сменяющих друг друга при смещении примерно на длину волны в любом направлении.

где  $\rho = \rho_1 - \rho_2$ ,  $\rho_+ = (\rho_1 + \rho_2)/2$ . Это соотношение эквивалентно (11.16), если под приведенной (к плоскости  $z' = 0$ ) интенсивностью понимать величину

$$J(\rho') = \frac{1}{4\pi^2} g_Q(\omega) \bar{n}_s(\rho'), \quad (12.22)$$

где

$$\bar{n}_s(\rho') = \int \bar{n}(\mathbf{R}) dz' = \int \bar{n}(\rho', z') dz'$$

— число излучателей, приходящихся на единицу площади в плоскости  $z = 0$ .

Теорема Ван-Циттерта—Цернике сохраняет силу и в случае движущихся источников, но спектральная плотность  $g_Q(\omega)$ , входящая в определение приведенной интенсивности (12.22), а также в формулы (12.20) и (12.21), должна быть усреднена по всем возможным значениям доплеровского смещения частоты, возникающего из-за движения источников. Иными словами, в формулах (12.20)—(12.22) нужно произвести замену

$$g_Q(\omega) \rightarrow \langle g_Q(\omega - \Omega_d) \rangle, \quad (12.23)$$

где угловые скобки обозначают усреднение по распределению доплеровского сдвига  $\Omega_d = \omega v_{||}/c$ , пропорционального лучевой скорости  $v_{||} = (\mathbf{m}, \mathbf{v})$ ,  $\mathbf{m} = (\mathbf{r} - \mathbf{R})/|\mathbf{r} - \mathbf{R}|$ .

Замена (12.23) справедлива при условии, что «длина свободного пробега» источника  $l_{\text{своб}}$  (т. е. длина, на которой заметно меняется скорость излучателя, или же расстояние, проходимое источником за время корреляции излучаемого сигнала) мала по сравнению с расстоянием  $|\mathbf{r} - \mathbf{R}|$  до точки наблюдения:  $l_{\text{своб}} \ll |\mathbf{r} - \mathbf{R}|$ . В подавляющем большинстве практически интересных случаев это условие выполнено.

В качестве конкретного выражения для спектральной плотности  $g_Q(\omega)$  укажем на результаты расчетов, проведенных в ч. I (задача 11 к гл. VI) для пуассоновского импульсного процесса, состоящего из следующих друг за другом отрезков экспоненциально затухающих синусоидальных пучков. Этот процесс может служить моделью излучения атомов, возбуждаемых ударами соседей. Окончательная формула для спектральной плотности (формула (1) в указанной задаче) имеет вид

$$g_Q(\omega) = \frac{n_1(\gamma + n_1)}{\pi(2\gamma + n_1)} \overline{A^2} \left\langle \frac{1}{(\bar{\omega}_0 + \Omega_d - \omega)^2 + (\gamma + n_1)^2} \right\rangle, \quad (12.24)$$

где  $n_1$  — средняя частота соударений,  $\gamma$  — радиационная ширина линий излучения и одновременно показатель затухания импульсов вида  $Q(t) = A \exp\{-(\gamma + i\bar{\omega}_0)t + i\beta\}$ ,  $A$  — случайная амплитуда со средним квадратом  $\overline{A^2}$ ,  $\beta$  — случайная фаза, распределенная



равномерно в интервале  $(0, 2\pi)$ , а  $\bar{\omega}_0$  — частота излучения в системе отсчета, в которой осциллятор покоится. В случае максвелловского распределения скоростей частиц формула (12.24), в зависимости от соотношения между шириной линии излучения  $n_1 + \gamma$  и среднеквадратичным доплеровским уширением  $\sigma_\omega = \sqrt{\bar{\omega}_d^2}$ , дает либо гауссову (при  $\sigma_\omega \gg n_1 + \gamma$ ), либо лоренцеву (при  $\sigma_\omega \ll n_1 + \gamma$ ) форму линии излучения.

Теорема Ван-Циттерта — Цернике допускает обобщение еще в одном направлении. Спектральная плотность (12.24) соответствует модели, в которой излучают атомы только одного сорта и только на одной частоте. Между тем в реальных нагретых газах могут содержаться атомы многих сортов, излучающие на многих частотах; в плотном облаке многие световые импульсы поглощаются другими атомами, не достигнув наблюдателя, и т. д. Все эти эффекты можно учесть в теореме Ван-Циттерта — Цернике, если под  $J(\rho)$  понимать сумму приведенных интенсивностей, отвечающих различным линиям излучения:

$$J(\rho) = \sum_m J_m(\rho) = \frac{1}{4\pi^2} \sum_m g^{(m)}(\omega) \bar{n}_s^{(m)}(\rho).$$

Спектр каждой из этих линий излучения характеризуется собственной частотой излучателя  $\bar{\omega}_0^{(m)}$ , силой осциллятора  $\bar{A}_m^2$ , радиационной шириной  $\gamma_m$  и средним числом столкновений  $n_1^{(m)}$ . Величина  $\bar{n}_s^{(m)}(\rho)$  определяет тогда эффективную поверхностную концентрацию излучателей, импульсы от которых в состоянии выйти из объема  $V$ . Разумеется, если излучатели распределены не в объеме, а на поверхности, то при расчете следует с самого начала ввести вместо объемной концентрации  $\bar{n}(\mathbf{R})$  поверхностную концентрацию  $\bar{n}_s(\rho)$  (см., например, [39]).

### Задачи

1. Выразить среднюю плотность звуковой энергии в жидкости  $W = \frac{1}{2}\rho v^2 + \frac{1}{2}\beta p^2$  и средний вектор Умова (плотность потока энергии)  $\mathcal{S} = p\mathbf{v}$  через функцию когерентности потенциала скорости и через лучевую интенсивность.

Решение. Выразив давление  $p$  и скорость  $\mathbf{v}$  через потенциал скорости  $\Phi(t, \mathbf{r})$ :

$$p = -\rho \frac{\partial \Phi}{\partial t}, \quad \mathbf{v} = \nabla \Phi,$$

имеем

$$W = \frac{\rho}{2} \left[ (\nabla \Phi)^2 + \frac{1}{c^2} \left( \frac{\partial \Phi}{\partial t} \right)^2 \right], \quad \mathcal{S} = -\rho \frac{\partial \Phi}{\partial t} \nabla \Phi,$$

где  $c = (\beta\rho)^{-1/2}$  — скорость звука,  $\rho$  — средняя плотность,  $\beta$  — сжимаемость жидкости. Если  $u(t, \mathbf{r})$  — комплексный потенциал скорости (аналитический сигнал), отвечающий вещественному потенциалу  $\Phi(t, \mathbf{r})$ :  $\Phi = \text{Re } u$ , то  $W$  и  $\mathcal{S}$

записываются через  $u$  следующим образом:

$$W = \frac{\rho}{4} \left\{ |\nabla u|^2 + \frac{1}{c^2} \left| \frac{\partial u}{\partial t} \right|^2 + \operatorname{Re} \left[ (\nabla u)^2 + \frac{1}{c^2} \left( \frac{\partial u}{\partial t} \right)^2 \right] \right\},$$

$$\mathcal{S} = -\frac{\rho}{2} \operatorname{Re} \left( \frac{\partial u}{\partial t} \nabla u^* + \frac{\partial u}{\partial t} \nabla u \right).$$

Усреднение этих выражений дает

$$\langle W \rangle = \frac{\rho}{4} \left\{ \left( \nabla_1 \nabla_2 + \frac{1}{c^2} \frac{\partial^2}{\partial t_1 \partial t_2} \right) [\Gamma_u(1, 2) + \operatorname{Re} \tilde{\Gamma}_u(1, 2)] \right\}_{2 \rightarrow 1},$$

$$\langle \mathcal{S} \rangle = -\frac{\rho}{2} \operatorname{Re} \left\{ \frac{\partial}{\partial t_1} \nabla_2 [\Gamma_u(1, 2) + \tilde{\Gamma}_u(1, 2)] \right\}_{2 \rightarrow 1}. \quad (1)$$

Символом  $2 \rightarrow 1$  здесь обозначен предельный переход  $t_2 \rightarrow t_1$ ,  $\mathbf{r}_2 \rightarrow \mathbf{r}_1$ , который следует делать после вычисления временных и пространственных производных. Заметим, что закон сохранения  $\operatorname{div} \mathcal{S} + \partial W / \partial t = 0$ , который должен выполняться и для средних значений  $\langle W \rangle$  и  $\langle \mathcal{S} \rangle$ , при подстановке в него выражений (1) можно рассматривать как своеобразное уравнение сохранения для функций  $\Gamma_u$  и  $\tilde{\Gamma}_u$ .

Если поле  $u$  стационарно, то  $\bar{u} = 0$ ,  $\tilde{\Gamma} = 0$ , и тогда

$$\langle W \rangle = \frac{\rho}{4} \left[ \left( \nabla_1 \nabla_2 + \frac{1}{c^2} \frac{\partial^2}{\partial t_1 \partial t_2} \right) \Psi_u(1, 2) \right]_{2 \rightarrow 1},$$

$$\langle \mathcal{S} \rangle = -\frac{\rho}{2} \operatorname{Re} \left( \frac{\partial}{\partial t_1} \nabla_2 \Psi_u(1, 2) \right)_{2 \rightarrow 1}.$$

Если поле  $u$  к тому же и однородно, то его корреляционная функция  $\Psi_u(1, 2)$  имеет вид (9.27), что позволяет выразить  $\langle W \rangle$  и  $\langle \mathcal{S} \rangle$  через лучевую интенсивность:

$$\langle W \rangle = \frac{\rho}{2} \int_0^\infty k^2 d\omega \oint \mathcal{I}(\omega, \mathbf{n}) d\Omega(\mathbf{n}),$$

$$\langle \mathcal{S} \rangle = \frac{\rho}{2} \int_0^\infty k^2 d\omega \oint \mathbf{n} \mathcal{I}(\omega, \mathbf{n}) d\Omega(\mathbf{n}).$$

Для изотропного случайного звукового поля, очевидно,  $\langle \mathcal{S} \rangle = 0$ . Для квази-монохроматического звукового поля (спектр сосредоточен вблизи частоты  $\omega_0$ ) множитель  $k^2$  в написанных выражениях можно вынести за знак интеграла со значением  $k_0^2 = \omega_0^2 / c^2$ .

2. Найти общий вид функции корреляции изотропного, однородного и стационарного скалярного волнового поля.

Решение. Лучевая интенсивность  $\mathcal{I}(\omega, \mathbf{n})$  для изотропного поля не зависит от направления:  $\mathcal{I}(\omega, \mathbf{n}) = \mathcal{P}(\omega)$ , так что корреляционная функция (9.27) выражается однократным интегралом:

$$\Psi_u(t, \mathbf{r}) = 4\pi \int_0^\infty \mathcal{P}(\omega) \frac{\sin kr}{kr} d\omega,$$

поскольку

$$\oint e^{ik\mathbf{n}\mathbf{r}} d\Omega(\mathbf{n}) = 4\pi \frac{\sin kr}{kr}.$$

3. Оценить продольный и поперечный радиусы корреляции волнового поля, лучевая интенсивность которого  $\mathcal{J}(\omega, \mathbf{n})$  сосредоточена в узком конусе направлений с раствором  $\theta \ll 1$ .

Решение. Пусть ось  $z$  отвечает направлению, по которому лучевая интенсивность максимальна. Записав единичный вектор  $\mathbf{n}$  в виде

$$\mathbf{n} = \mathbf{l}_z n_z + \mathbf{n}_\perp,$$

при  $n_\perp \ll \theta \ll 1$  имеем  $n_z = \sqrt{1 - n_\perp^2} \approx 1 - n_\perp^2/2$ , так что  $(\mathbf{n}\mathbf{r}) \approx z - 1/2 n_\perp^2 z + \mathbf{n}_\perp \cdot \mathbf{p}$ ,  $\mathcal{J}(\omega, \mathbf{n}) \approx \mathcal{J}(\omega, \mathbf{n}_\perp, 1)$  и корреляционная функция (9.27) принимает вид

$$\Psi(t, \mathbf{r}) = \int_0^\infty d\omega \exp(ikz - i\omega t) \oint \mathcal{J}(\omega, \mathbf{n}_\perp, 1) \exp\left(ik\mathbf{n}_\perp \cdot \mathbf{p} - \frac{i}{2} k n_\perp^2 z\right) d\mathbf{o}(\mathbf{n}), \quad (1)$$

где  $\mathbf{p} = (x, y)$ . Поперечный ( $l_\perp$ ) и продольный ( $l_\parallel$ ) радиусы корреляции оцениваются из условий  $|k\mathbf{n}_\perp \cdot \mathbf{p}| \ll 1$ ,  $|k n_\perp^2 z| \ll 1$ , если подставить в них  $n_\perp \approx \theta$ . В результате находим

$$l_\perp \sim \frac{1}{k\theta} \sim \frac{\lambda}{\theta}, \quad l_\parallel \sim \frac{1}{k\theta^2} \sim \frac{\lambda}{\theta^2}. \quad (2)$$

4. Установить связь между лучевой интенсивностью и двумерным спектром (9.29) в случае узкого углового спектра [40].

Решение. В узком конусе  $|n_\perp| \ll \theta \ll 1$  выполняется приближенное соотношение  $d\mathbf{o}(\mathbf{n}) \approx dn_x dn_y$ , которое можно представить в эквивалентной форме  $d\mathbf{o}(\mathbf{n}) = k^{-2} d\kappa_x d\kappa_y = k^{-2} d^2\kappa_\perp$ , если ввести двумерный волновой вектор  $\kappa_\perp = k\mathbf{n}_\perp = (k n_x, k n_y)$ . Формула (1) предыдущей задачи принимает тогда вид двумерного спектрального разложения:

$$\Psi(t, \mathbf{r}) = \int_0^\infty d\omega \exp(ikz - i\omega t) \int_{-\infty}^\infty k^{-2} \mathcal{J}(\omega, \mathbf{n}_\perp, 1) \exp\left(i\kappa_\perp \cdot \mathbf{p} - \frac{i\kappa_\perp^2 z}{2k}\right) d^2\kappa_\perp.$$

Сравнение этого разложения с разложением

$$\Psi(t, \mathbf{r}) = \int_0^\infty d\omega \int_{-\infty}^\infty F(\omega; \kappa_\perp, z) \exp(i\kappa_\perp \cdot \mathbf{p} - i\omega t) d^2\kappa_\perp,$$

обратным (9.29), показывает, что

$$F(\omega; \kappa_\perp, z) = k^{-2} \mathcal{J}(\omega, \mathbf{n}_\perp, 1) \exp\left(ikz - \frac{i\kappa_\perp^2 z}{2k}\right).$$

5. Выразить среднюю плотность электромагнитной энергии и средний вектор Пойнтинга в вакууме через матрицу лучевых интенсивностей  $\mathcal{J}_{\alpha\beta}$ , которая вводится аналогично (9.26):

$$G_{\alpha\beta}^{(E)}(\omega, \kappa) = \mathcal{J}_{\alpha\beta}(\omega, \mathbf{n}) k^{-2} \delta(\kappa - k), \quad \mathbf{n} = \frac{\kappa}{\kappa}, \quad k \equiv \frac{\omega}{c};$$

при этом, подобно (9.27),

$$\Psi_{\alpha\beta}^{(E)}(t, \mathbf{r}) = \int_0^\infty d\omega \oint \mathcal{J}_{\alpha\beta}(\omega, \mathbf{n}) e^{i(k\mathbf{n}\mathbf{r} - \omega t)} d\mathbf{o}(\mathbf{n}). \quad (1)$$

Здесь  $G_{\alpha\beta}^{(E)}$  — матрица спектральных плотностей, представляющая собой преобразование Фурье корреляционной матрицы свободного статистически однородного электромагнитного поля

$$\Psi_{\alpha\beta}^{(E)}(t, \mathbf{r}) = \langle \bar{E}_{\alpha}(t_1, \mathbf{r}_1) \bar{E}_{\beta}^*(t_2, \mathbf{r}_2) \rangle.$$

Решение. Рассматривая случай стационарного и однородного электромагнитного поля с нулевым средним значением и учитывая, что в вакууме  $|\mathbf{H}|^2 = |\mathbf{E}|^2$ , имеем

$$\langle W \rangle = \frac{1}{16\pi} (\langle |\mathbf{E}|^2 \rangle + \langle |\mathbf{H}|^2 \rangle) = \frac{1}{8\pi} \langle |\mathbf{E}|^2 \rangle.$$

Но при  $\bar{\mathbf{E}} = 0$  в силу (1)

$$\langle |\mathbf{E}|^2 \rangle = \sum_{\alpha=1}^3 \Psi_{\alpha\alpha}^{(E)}(0, 0) = \int_0^{\infty} d\omega \oint \mathcal{J}(\omega, \mathbf{n}) d\mathbf{n},$$

где введена лучевая интенсивность

$$\mathcal{J}(\omega, \mathbf{n}) = \sum_{\alpha=1}^3 \mathcal{J}_{\alpha\alpha}(\omega, \mathbf{n}). \quad (2)$$

В результате имеем

$$\langle W \rangle = \frac{1}{8\pi} \int_0^{\infty} d\omega \oint \mathcal{J}(\omega, \mathbf{n}) d\mathbf{n}.$$

Среднее значение вектора Пойнтинга

$$\langle \mathcal{S} \rangle = \frac{c}{8\pi} \operatorname{Re} \langle [\mathbf{E}\mathbf{H}^*] \rangle$$

можно получить, используя спектральные разложения для  $\mathbf{E}$  и  $\mathbf{H}$  и соотношение  $\mathbf{H}_{\omega\mathbf{k}} = [\mathbf{n}\mathbf{E}_{\omega\mathbf{k}}]$ :

$$\begin{aligned} \langle \mathcal{S} \rangle &= \frac{c}{8\pi} \operatorname{Re} \int_{-\infty}^{\infty} d^3\mathbf{k}' d^3\mathbf{k}'' \int_0^{\infty} d\omega' d\omega'' \times \\ &\times \langle [\mathbf{E}_{\omega'\mathbf{k}'}^* [\mathbf{n}''\mathbf{E}_{\omega''\mathbf{k}''}^*]] \rangle \exp[i(\mathbf{k}' - \mathbf{k}'')\mathbf{r} - i(\omega' - \omega'')t], \end{aligned} \quad (3)$$

где  $\mathbf{n}'' = \mathbf{k}''/k''$ . Раскрывая двойное векторное произведение, получаем выражение

$$\langle \dots \rangle = \mathbf{n}'' \langle \mathbf{E}_{\omega'\mathbf{k}'} \mathbf{E}_{\omega''\mathbf{k}''}^* \rangle - \langle \mathbf{E}_{\omega''\mathbf{k}''}^* (\mathbf{n}'' \mathbf{E}_{\omega'\mathbf{k}'}^*) \rangle.$$

Вследствие дельта-корреляции спектральных амплитуд вектор  $\mathbf{n}''$  можно заменить на  $\mathbf{n}' = \mathbf{k}'/k'$ , но в силу поперечности электромагнитного поля  $(\mathbf{n}' \mathbf{E}_{\omega'\mathbf{k}'}^*) = 0$ .

В результате, с учетом (2), имеем

$$\begin{aligned} \langle \dots \rangle &= \sum_{\alpha=1}^3 \mathbf{n}' G_{\alpha\alpha}(\omega', \mathbf{k}') \delta(\omega' - \omega'') \delta(\mathbf{k}' - \mathbf{k}'') = \\ &= \mathbf{n}' \mathcal{J}(\omega, \mathbf{n}) k^{-2} \delta(\omega' - \omega'') \delta(\mathbf{k}' - \mathbf{k}'') \delta(\mathbf{k}' - \mathbf{k}). \end{aligned}$$

Подставляя это выражение в (3), получаем

$$\langle \mathcal{S} \rangle = \frac{c}{8\pi} \int_0^{\infty} d\omega \oint \mathbf{n} \mathcal{J}(\omega, \mathbf{n}) d\mathbf{n}. \quad (4)$$

Легко видеть, что электромагнитные формулы для  $\langle W \rangle$  и  $\langle \mathcal{P} \rangle$  отличаются от акустических выражений (см. задачу 1) только множителем  $1/8\pi$  вместо  $\rho k^2/2$ .

Из (4) следует, что величина  $\frac{c}{8\pi} \mathbf{n} \mathcal{I}(\omega, \mathbf{n})$  имеет смысл потока энергии через единичную площадку в единичный телесный угол в расчете на единичный частотный интервал. Именно так вводится лучевая интенсивность (яркость) в феноменологической теории переноса излучения [41, 42].

6. Построить корреляционную матрицу  $\Psi_{\alpha\beta}^{(E)}$  для изотропного, однородного и стационарного электромагнитного поля.

Решение. Входящая в формулу (1) предыдущей задачи матрица  $\mathcal{I}_{\alpha\beta}$  в случае изотропного поля должна иметь вид

$$\mathcal{I}_{\alpha\beta}(\omega, \mathbf{n}) = \mathcal{P}(\omega) \delta_{\alpha\beta} + Q(\omega) n_{\alpha} n_{\beta}.$$

Кроме того, в силу поперечности электромагнитных волн в свободном пространстве матрица  $\mathcal{I}_{\alpha\beta}$  должна удовлетворять условиям

$$n_{\alpha} \mathcal{I}_{\alpha\beta} = n_{\beta} \mathcal{I}_{\alpha\beta} = 0.$$

Это возможно только при  $Q = -\mathcal{P}$ , и тогда

$$\begin{aligned} \Psi_{\alpha\beta}^{(E)}(t, \mathbf{r}) &= \int_0^{\infty} d\omega \mathcal{P}(\omega) \oint e^{i\mathbf{k}\mathbf{n}\mathbf{r} - i\omega t} (\delta_{\alpha\beta} - n_{\alpha} n_{\beta}) d\mathbf{n} = \\ &= \int_0^{\infty} d\omega \mathcal{P}(\omega) e^{-i\omega t} \left( \delta_{\alpha\beta} - \frac{1}{(ik)^2} \frac{\partial^2}{\partial r_{\alpha} \partial r_{\beta}} \right) \oint e^{i\mathbf{k}\mathbf{n}\mathbf{r}} d\mathbf{n} = \\ &= \int_0^{\infty} d\omega \mathcal{P}(\omega) e^{-i\omega t} k^{-2} \left\{ \left[ k^2 s + \frac{s'}{r} \right] \delta_{\alpha\beta} + \left[ s'' - \frac{s'}{r} \right] \frac{r_{\alpha} r_{\beta}}{r^2} \right\}, \end{aligned}$$

где  $s = 4\pi \sin kr / kr$ , а штрихом обозначено дифференцирование по  $r$ . Лучевая интенсивность  $\mathcal{I} = \sum \mathcal{I}_{\alpha\alpha}$  в данном случае равна  $2\mathcal{P}(\omega)$  и, следовательно, не зависит от направления; при этом, в соответствии с результатами задачи 5,

$$\langle \mathcal{P} \rangle = 0, \text{ а } \langle W \rangle = \int_0^{\infty} \mathcal{P}(\omega) d\omega.$$

7. Вычислить индекс мерцаний в дальней зоне для амплитудно-фазового экрана с логарифмически нормальным законом распределения вероятностей поля.

Решение. Пусть  $v = e^{\chi + iS}$ , где фаза  $S$  и уровень  $\chi$  распределены по нормальному закону, так что  $v$  — логарифмически нормальное поле. Предположим, что  $\bar{S} = 0$  и  $\langle S\chi \rangle = 0$ . В силу нормировки (10.23)  $\bar{\chi} = -\sigma_{\chi}^2$ . В результате

$$v_0 = \langle e^{\chi + iS} \rangle = e^{\bar{\chi} + 1/2 \langle (\bar{\chi} + iS)^2 \rangle} = e^{-1/2 (\sigma_S^2 + \sigma_{\chi}^2)}.$$

Следовательно, в дальней зоне

$$\beta^2 = \beta_{\infty}^2 = 1 - |v_0|^4 = 1 - \exp \{ -2 (\sigma_S^2 + \sigma_{\chi}^2) \}.$$

В случае чисто амплитудных флуктуаций ( $\sigma_S^2 = 0$ ) квадрат индекса мерцаний в дальней зоне равен  $\beta_\infty^2 = 1 - e^{-2\sigma_0^2}$ , тогда как на самом экране  $\beta^2 = \beta_0^2 = e^{4\sigma_0^2} - 1$ . Очевидно,  $\beta_0^2 > 1 > \beta_\infty^2$ , т. е. амплитудные флуктуации за экраном сглаживаются.

8. Пусть поле в плоскости экрана *вещественно* и распределено по нормальному закону. Показать, что квадрат индекса мерцаний в дальней зоне вдвое меньше, чем на экране.

Решение. В дальней зоне, согласно (10.38),  $\beta_\infty^2 = 1 - |v_0|^4$ . На экране же

$$\beta^2 = \beta_0^2 = \langle v^4 \rangle - 1 = v_0^4 + 6v_0^2 \langle \tilde{v}^2 \rangle + \langle \tilde{v}^4 \rangle.$$

Но для вещественного нормального случайного поля  $\langle \tilde{v}^4 \rangle - 3\langle \tilde{v}^2 \rangle^2 = 3\sigma_v^2$ . Поэтому, с учетом нормировки  $\sigma_v^2 + v_0^2 = 1$ , имеем

$$\beta_0^2 = v_0^4 + 6v_0^2(1 - v_0^2) + 3(1 - v_0^2)^2 = 2(1 - v_0^4) = 2\beta_\infty^2,$$

т. е.  $\beta_\infty^2 = 1/2 \beta_0^2$ . В этой и предыдущей задачах  $\beta_\infty^2 < \beta_0^2$ , потому что поле  $v$  порождает за экраном помимо амплитудных еще и фазовые флуктуации, которые, очевидно, не влияют на величину флуктуаций интенсивности.

9. Вывести из френелевского приближения (9.17) соотношение (10.13), согласно которому при падении плоской волны корреляционная функция за бесконечным экраном  $\psi_\perp(\rho)$  совпадает с функцией корреляции граничного поля  $\psi_v(\rho)$ .

Решение. Во френелевском приближении

$$\psi_\perp(\rho_1, \rho_2, z) =$$

$$= \left( \frac{k}{2\pi z} \right)^2 \iint_{-\infty}^{\infty} \psi_v(\rho' - \rho'') \exp \left\{ \frac{ik}{2z} [(\rho_1 - \rho')^2 - (\rho_2 - \rho'')^2] \right\} d^2\rho' d^2\rho''.$$

Вводя обозначения  $\rho = \rho_1 - \rho_2$ ,  $\rho_+ = (\rho_1 + \rho_2)/2$ ,  $\xi = (\rho' - \rho'')$ ,  $\eta = (\rho' + \rho'')/2$ , получаем выражение

$$\psi_\perp(\rho, \rho_+, z) = \left( \frac{k}{2\pi z} \right)^2 \iint_{-\infty}^{\infty} \psi_v(\xi) \exp \left\{ \frac{ik}{z} [\rho\rho_+ + \xi\eta - \eta\rho - \xi\rho_+] \right\} d^2\xi d^2\eta.$$

Интеграл по  $\eta$  здесь вычисляется точно и равен  $\left( \frac{2\pi z}{k} \right)^2 \delta(\xi - \rho)$ , что и приводит к соотношению  $\psi_\perp(\rho, z) = \psi_v(\rho)$ .

10. Найти закон изменения среднего значения и функции корреляции сферической волны, прошедшей через неограниченный статистически однородный хаотический экран.

Решения. Пусть слева от плоскости  $z=0$ , в которой расположен экран, находится источник, излучающий сферическую волну

$$\exp \{ ik |\mathbf{r} - \mathbf{r}_0| \} / |\mathbf{r} - \mathbf{r}_0|,$$

где  $\mathbf{r}_0 = (0, -z_0)$  (рис. 10). Если  $f(\rho)$  — комплексная функция пропускающая экрана, то граничное поле в плоскости  $z=0$  равно

$$v(\rho') = f(\rho') \frac{\exp \left( ik \sqrt{z_0^2 + \rho'^2} \right)}{\sqrt{z_0^2 + \rho'^2}}.$$

Подставим это выражение для поля в формулу Грина (9.12) и, зафиксировав точку наблюдения  $\mathbf{r} = (\rho, z)$ , разложим показатель экспоненты в ряд по степеням отклонения  $\delta = \rho' - \rho_{\text{ст}}$  переменной  $\rho'$  от стационарной точки  $\rho'_{\text{ст}} = \rho z / (z + z_0)$ , лежащей на луче, соединяющем точку наблюдения с источником. Все предэкспоненциальные множители, кроме  $f(\rho)$ , заменим стационарными значениями, т. е. значениями при  $\rho' = \rho'_{\text{ст}}$ .

Получившуюся приближенную формулу

$$u(\rho, z) \approx \frac{ke^{ikR_0}(z+z_0)^3}{2\pi i R_0^3 z z_0} \int_{-\infty}^{\infty} f(\rho_{\text{ст}} + \delta) \exp\left(\frac{ik}{2L} \left[\delta^2 - \frac{(\delta\rho)^2}{R_0^2}\right]\right) d^2\delta \quad (1)$$

можно назвать френелевским приближением для сферической волны. Здесь

$$R = |\mathbf{r} - \mathbf{r}_0| = \sqrt{(z+z_0)^2 + \rho^2}, \quad L = R_0 z z_0 / (z + z_0)^2.$$

Из (1) находим среднее поле

$$\bar{u}(\mathbf{r}) = \bar{f} \frac{e^{ikR_0}}{R_0},$$

представляющее собой сферическую волну с амплитудой  $\bar{f}$ , и функцию корреляции поля в плоскости  $z = \text{const}$ :

$$\psi_u(\rho_1, \rho_2, z) = \psi_u(\rho, \rho_+, z) = \frac{\exp\left(ik \frac{\rho\rho_+}{R_0}\right)}{R_0^2} \psi_f\left(\rho \frac{z}{z+z_0}\right),$$

где  $\rho = \rho_1 - \rho_2$ ,  $\rho_+ = (\rho_1 + \rho_2)/2$ . Аналогичное выражение получается и для функции когерентности.

В отличие от рассмотренного в § 10 случая плоской волны, функция корреляции сферической волны, прошедшей через экран, не сохраняется. В частности, радиус корреляции поля в плоскости  $z = \text{const}$  растет при удалении от экрана пропорционально расстоянию от источника  $l_u = l_f \frac{z+z_0}{z_0}$ , где  $l_f$  — радиус корреляции функции пропускания  $f$ . Роль волнового параметра в данной задаче играет величина  $D_{\text{сф}} = 2L/kl_f^2$ , которая при  $z \rightarrow \infty$  стремится к конечному пределу  $D_{\text{max}} = 2z_0/kl_f^2$ . Легко понять, что

при  $D_{\text{max}} \ll 1$ , когда радиус первой зоны Френеля  $\sqrt{\lambda z z_0 / (z + z_0)}$  при любых  $z$  меньше размера неоднородностей на экране  $l_f$ , поле сферической волны за экраном не нормализуется даже на бесконечности.

11. Найти распределение интенсивности поля в дальней зоне ограниченного хаотического экрана.

Решение. Поле в зоне Фраунгофера определяется выражением (11.11), в котором под  $v$  нужно понимать граничное поле  $v(\rho) = f(\rho) u_n$ , возникающее при падении плоской волны  $u_n = e^{ikz}$  на хаотический экран с функцией пропускания  $f(\rho)$ . Распределение интенсивности в дальней зоне можно вычислить из (11.11) так же, как это было сделано в § 12 для поля (12.3), возбужденного антенной с флуктуирующими токами. Учитывая, что выражение (12.3) переходит в (11.11) при замене  $q \rightarrow -2ikv$ , по аналогии с распределе-

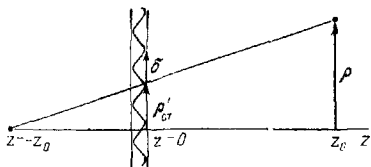


Рис. 10.

нием интенсивности (12.11) можно написать

$$\bar{I}_u(\rho_+, z) = \left( \frac{k}{2\pi z} \right)^2 \left[ |\bar{v}|^2 S^2 \left| \mathcal{F}_M \left( \frac{k\rho_+}{z} \right) \right|^2 + \sigma_v^2 S_v S \mathcal{F}_v \left( \frac{k\rho_+}{z} \right) \right], \quad (1)$$

где величины  $S_v$  и  $\mathcal{F}_v$  вводятся по аналогии с (12.8).

Заметим, что подобное же распределение интенсивности (с заменой  $z \rightarrow F$ ) получится в фокальной плоскости линзы, поставленной непосредственно после

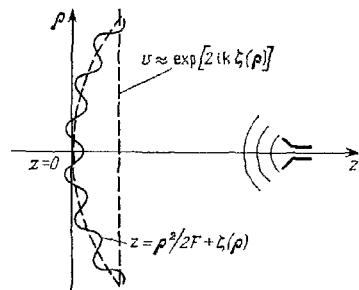


Рис. 11.

ограниченного хаотического экрана. В обоих случаях при переходе от слабых флуктуаций ( $\sigma_v \ll 1$ ) к сильным ( $\sigma_v \gg 1$ ) распределение интенсивности трансформируется так же, как на рис. 8.

12. Найдите средний коэффициент направленного действия параболитической зеркальной антенны с пологими неровностями.

Решение. Пусть  $z = \rho^2/2F$  — уравнение невозмущенного параболоида, а  $z = \rho^2/2F + \xi(\rho)$  — уравнение зеркала при наличии неровностей  $\xi(\rho)$ . Облучатель, помещенный в фокусе параболоида (рис. 11), посылает на зеркало сферическую волну, которая после отражения превращается в искаженную

плоскую волну. При плавных (в масштабе  $\lambda$ ) неровностях зеркала искажение поля в апертуре можно учесть введением фазового множителя  $\exp[iS(\rho)] = \exp[2ik\xi(\rho)]$  ( $2k\xi$  — дополнительный набег фазы, обусловленный неровностями). Тем самым задача сведена к расчету дифракции поля с граничным значением на апертуре  $v(\rho) = \exp[2ik\xi(\rho)]$ .

Если высота неровностей  $\xi(\rho)$  распределена по нормальному закону, то

$$\bar{v} = \exp(-2k^2\sigma_\xi^2),$$

$$\sigma_v^2 = 1 - |\bar{v}|^2 = 1 - e^{-4k^2\sigma_\xi^2};$$

при этом для малых неровностей

$$\sigma_v^2 \approx 4k^2\sigma_\xi^2 \ll 1, \quad l_v \sim l_\xi \quad (k\sigma_\xi \ll 1),$$

а для больших неровностей

$$\sigma_v^2 \approx 1, \quad l_v \sim l_\xi/k\sigma_\xi \ll l_\xi \quad (k\sigma_\xi \gg 1),$$

где  $l_\xi$  — радиус корреляции неровностей, а величина  $4k^2\sigma_\xi^2$  играет ту же роль, что и дисперсия фазы  $\sigma_S^2$  в теории фазового экрана.

Используя для расчета к. н. д. формулу (12.14), в которой, с учетом предыдущей задачи, можно заменить  $S_q$ ,  $\sigma_q^2$ ,  $|\bar{q}|^2$  и  $Q^2 = \sigma_q^2 + |\bar{q}|^2$  соответственно на  $S_v$ ,  $\sigma_v^2$ ,  $|\bar{v}|^2$  и  $\sigma_v^2 + |\bar{v}|^2 = 1$ , получаем

$$\bar{G} = G_0 \left\{ 1 - \left( 1 - \frac{S_v}{S} \right) \sigma_v^2 \right\} = G_0 \left\{ \frac{S_v}{S} + \left( 1 - \frac{S_v}{S} \right) e^{-4k^2\sigma_\xi^2} \right\}$$



(здесь, кроме того, учтено, что  $S_A = S$  при  $\bar{v} = \text{const}$ ). В предельных случаях имеем

$$\begin{aligned}\bar{G} &= G_0 (1 - \sigma_v^2) \approx G_0 (1 - 4k^2 \sigma_v^2) \quad (k\sigma_v \ll 1), \\ \bar{G} &\approx G_0 \left( \frac{S_v}{S} + e^{-4k^2 \sigma_v^2} \right) \quad (k\sigma_v \gg 1),\end{aligned}\quad (1)$$

или, если принять во внимание оценку  $S_v \sim l_v^2 \sim l_\xi^2 / (k\sigma_\xi)^2$ ,

$$\bar{G} \approx G_0 \left( e^{-4k^2 \sigma_\xi^2} + \frac{l_\xi^2}{S k^2 \sigma_\xi^2} \right) \quad (k\sigma_\xi \gg 1). \quad (2)$$

Качественный ход зависимости нормированного среднего к. н. д.  $\bar{G}/G_0$  от среднеквадратичного набег фазы  $k\sigma_\xi$  показан на рис. 12. Начальный участок кривой отвечает первой формуле (1), а конечный — формуле (2). Из рисунка видно, что заметное уменьшение к. н. д. начинается со значений  $k\sigma_\xi \sim 1$ , т. е. при  $\sigma_\xi \sim 1/k = \lambda/2\pi$ , что согласуется с инженерным критерием гладкости зеркал  $\sigma_\xi < \lambda/8$ . Для больших зеркальных антенн современных радиотелескопов величина  $\sigma_\xi$  может быть доведена примерно до сантиметра. В соответствии с критерием  $\sigma_\xi < \lambda/8$  резкое снижение эффективности таких антенн происходит при переходе от сантиметрового к миллиметровому диапазону волн.

13. Оценить продольный радиус корреляции поля, созданного облаком статистически независимых источников.

Решение. Расположим точки наблюдения  $r_1$  и  $r_2$  на оси  $z$ :  $r_1 = (0, 0, z + \xi)$ ,  $r_2 = (0, 0, z)$ . Для входящей в формулу (12.20) разности хода  $s$  имеем

$$s = |r_1 - R| - |r_2 - R| = \sqrt{(z + \xi - z')^2 + \rho'^2} - \sqrt{(z - z')^2 + \rho'^2}, \quad (1)$$

где переменные интегрирования  $\rho'$  и  $z'$  меняются в пределах объема  $V$ , занятого источниками:  $|\rho'| \leq L$ ,  $|z'| \leq L$ . При анализе продольной корреляции рассмотрим три характерные зоны.

В ближней зоне ( $z \leq L$ ) можно ожидать малости  $l_\parallel$  по сравнению с  $L$ , что позволяет разложить (1) в степенной ряд по  $\xi$  и ограничиться только линейным членом:

$$s \approx \xi |z - z'| / \sqrt{|z - z'|^2 + \rho'^2}. \quad (2)$$

При всех значениях  $\rho'$  и  $z'$  множитель при  $\xi$  не превышает единицы. Поэтому присутствие множителя  $\exp(iks)$  под знаком интеграла в (12.20) проявится лишь при  $k\xi \gg 1$ . Следовательно, продольный радиус корреляции может быть оценен из условия  $k\xi \sim 1$ , т. е.

$$l_\parallel \sim \frac{1}{k} \sim \lambda \quad (z \leq L). \quad (3)$$

В промежуточной зоне ( $L \ll z \ll kL^2$ ) отношения  $z'/z$  и  $\rho'/z$  малы по сравнению с единицей, в результате чего выражение (2) упрощается и

принимает вид

$$s \approx \zeta \left( 1 - \frac{\rho'^2}{2z^2} \right). \quad (4)$$

Первое слагаемое в (4) не зависит от переменных интегрирования и его можно не принимать во внимание при рассмотрении корреляции, так как  $|e^{ik\zeta}| = 1$ . Второе слагаемое по модулю не превышает величины  $\zeta L^2/2z^2$ . Из условия  $k\zeta L^2/z^2 \sim 1$  получаем оценку

$$I_1 \sim \lambda \frac{z^2}{L^2} \quad (L \ll z \ll kL^2). \quad (5)$$

Эта оценка аналогична формуле (2) из задачи 3 для продольного радиуса корреляции однородного случайного поля с шириной углового спектра  $\theta \sim a/z$ . Можно сказать, что в данном случае ширина углового спектра  $\theta$  определяется видимым угловым размером облака источников  $\gamma \sim L/z$ .

В *дальней* зоне ( $z \gg kL^2$ ) величина  $ks$  для всех  $\rho'$  и  $z'$  отличается от  $k\zeta$  не более чем на  $\pi$ . Поэтому в дальней зоне  $e^{iks} \approx e^{ik\zeta}$ , а продольный радиус корреляции увеличивается до бесконечности:

$$I_1 = \infty \quad (z \gg kL^2), \quad (6)$$

что отвечает полностью сформировавшейся диаграмме излучения в зоне Фраунгофера.

## Глава III

### ТЕПЛОВОЕ ЭЛЕКТРОМАГНИТНОЕ ПОЛЕ

#### § 13. Предварительные замечания

Тепловое электромагнитное поле создается в результате хаотического теплового движения заряженных микрочастиц (электронов, ионов и т. п.), из которых построены тела. Уже отсюда ясно, что интенсивность теплового поля должна возрастать с повышением температуры, подобно тому как с ростом температуры убыстряется брауновское движение или усиливаются тепловые шумы в электрических цепях. Но в макроскопической электродинамике электромагнитное поле рассматривается не как сумма полей отдельных микрочастиц, а как поле макроскопических же источников, описываемых, например, объемными плотностями заряда и тока. Применительно к тепловому электромагнитному полю это означает, что его источниками являются *пространственно-временные флуктуации заряда и тока* в физически бесконечно малых элементах объема тел. Чаще всего масштабом малости таких объемов служит при спектральном описании поля интересующая нас длина волны, но аналогичную роль могут играть и размеры тел или их неоднородностей, расстояния между телами и т. п. Если физические и геометрические условия задачи допускают выделение таких физически бесконечно малых элементов объема, то можно пользоваться уравнениями макроскопической электродинамики, т. е. уравнениями Максвелла.

Мы уже рассматривали тепловые флуктуации электрических величин, но в *квазистационарной* области, т. е. для электрических систем, размеры которых  $l$  много меньше длины волны  $\lambda$ :  $l \ll \lambda$ <sup>1)</sup>. Состояние таких систем можно описать конечным числом *интегральных* функций времени — сил токов через сечения проводов, зарядов на сосредоточенных емкостях и т. п. При помощи теоремы Найквиста и обобщающей ее флуктуационно-диссипационной теоремы (ФДТ) мы получили в ч. I, § 54 выражения для вторых моментов *термодинамически равновесных* флуктуаций таких интегральных величин. Мы уже знакомы, таким

<sup>1)</sup> Не следует смешивать квазистационарность в этом электродинамическом смысле со *статистической* квазистационарностью случайных процессов.

образом, с корреляционной теорией электрических тепловых флуктуаций в квазистационарных цепях, т. е. в случае  $l \ll \lambda$ .

Но существует и совсем иная область явлений, для которых теория тепловых электромагнитных полей была развита уже давно, причем при прямо противоположных условиях, когда все характерные размеры тел  $l$  велики по сравнению с длиной волны:  $l \gg \lambda$ . Это классическая теория теплового излучения, созданная Г. Кирхгофом, Л. Больцманом и рядом других выдающихся физиков конца прошлого века. Ее завершило в 1900 г. открытие М. Планком закона, дающего спектральное распределение энергии равновесного теплового излучения, заключенного в полости достаточно больших размеров:  $l \gg \lambda$ . Другие законы этой теории (например, законы Кирхгофа) тоже опираются на асимптотическое описание электромагнитного поля в *приближении геометрической оптики*, пригодном только для достаточно коротких волн. Исторически это произошло потому, что классическая теория теплового излучения развивалась прежде всего для *оптического* диапазона.

Общая флуктуационная электродинамика полей теплового происхождения, основанная на уравнениях Максвелла (понимаемых, конечно, как *стохастические* уравнения), должна быть, очевидно, свободна от ограничений, касающихся соотношения между размерами тел  $l$  и длиной волны  $\lambda$ . Она должна охватывать, наряду с найквистовской теорией тепловых шумов в квазистационарных цепях ( $l \ll \lambda$ ) и классической теорией теплового излучения ( $l \gg \lambda$ ), также промежуточную область  $l \sim \lambda$ , в которой оба предельных подхода неприменимы. Такая общая теория тепловых электромагнитных полей была построена сравнительно недавно [1], и в данной главе излагаются ее основы. Для краткости мы будем далее называть случайное электромагнитное поле теплового происхождения просто *флуктуационным* полем.

Прежде чем перейти к систематическому изложению, остановимся на двух вопросах, которые легко могут возникнуть после сказанного выше. Один из них можно поставить следующим образом: почему в классической теории теплового излучения речь идет не вообще о тепловом флуктуационном поле, а именно об *излучении*, т. е. только о *волновом* поле?

Это вполне резонный вопрос, потому что флуктуации заряда и тока в каждом элементе объема тела порождают не только разбегающуюся электромагнитную волну, но и так называемое *ближнее* квазистационарное поле. Известно, однако, что оно убывает с расстоянием  $R$  от своего «точечного» источника гораздо быстрее, чем волновое поле, — не как  $1/R$ , а как  $1/R^2$  и  $1/R^3$ , простираясь лишь на расстояния порядка  $\lambda$ . Кроме того, это ближнее поле в среднем за период не создает потока энергии. Именно поэтому классическая теория теплового излучения, инте-

ресующаяся только *переносом* «лучистой» энергии и только достаточно короткими волнами ( $\lambda \ll l$ ), просто игнорирует ближнее флуктуационное поле. Она не может поэтому правильно описать флуктуационное поле, например, в полости, размеры которой невелики или тем более малы по сравнению с длиной волны. Равным образом она оставляет без внимания особенности теплового поля на малых расстояниях ( $l \lesssim \lambda$ ) от поверхности нагретого тела. Между тем нетрудно понять, в чем заключаются эти особенности и каково их происхождение.

Внутри нагретого тела, в каждом элементе его объема происходят непрерывные тепловые флуктуации плотностей заряда и тока, непрерывное электрическое «кишение», если воспользоваться образным выражением М. А. Леонтовича. Сумма *волновых* полей этих источников дает вне тела его тепловое электромагнитное излучение, описываемое законом Кирхгофа. Сумма же *ближних* полей (от приповерхностных элементов объема) создает *квазистационарное* тепловое поле, которое как бы *выстилает поверхность* любого тела. Это поле сосредоточено около поверхности, в слое толщины порядка  $\lambda$ , так что на расстояниях  $l \gg \lambda$  им можно пренебречь, но при  $l \rightarrow 0$  его напряженность *неограниченно* возрастает. В этом, конечно, нет никакого парадокса, так как на удалениях от поверхности порядка межатомных расстояний макроскопическое описание поля уже становится непригодным.

Хотя квазистационарное флуктуационное поле не участвует в создании потока энергии, его наличие проявляется вполне наблюдаемым образом, например, в молекулярных силах сцепления между телами (см. § 21).

Использование приближения геометрической оптики ограничивает классическую теорию теплового излучения не только тем, что она не учитывает квазистационарного флуктуационного поля, но и тем, что от нее ускользают *дифракционные* явления в самом излучении.

Другой законный вопрос касается того, представляет ли интерес область  $l \sim \lambda$ , т. е. тот случай, когда размеры тел сравнимы с длиной волны.

При своем зарождении радиотехника использовала метровые и даже сантиметровые волны, возникающие при колебательном искровом разряде. Весьма быстро, однако, она была вынуждена перейти к очень длинным (километровым) волнам, после чего — уже с электронными лампами в генераторах и усилителях — начался длительный процесс настоящего технического освоения все более коротких волн. Постепенно этот путь привел к дециметровому диапазону, а затем — к сантиметровым и миллиметровым волнам и, соответственно, к радиотехническим устройствам, у которых все (или некоторые) размеры сравнимы с длиной

волны, т. е. к не квазистационарным *распределенным волновым* системам. Такого рода устройства (волноводы, объемные резонаторы, рупорные и зеркальные антенны и т. п.) составили предмет так называемой электродинамики сверхвысоких частот (СВЧ). По мере повышения чувствительности и точности применяемых радиотехнических методов естественно приобрели остроту вопросы, касающиеся тепловых шумов и в распределенных системах. Таким образом, если говорить о нуждах радиотехники, то развитие общей флуктуационной электродинамики было стимулировано проблемой тепловых шумов в диапазоне СВЧ. Однако физические результаты флуктуационной электродинамики, конечно, не исчерпываются ее радиотехническими приложениями. Например, одним из основных ее следствий является обобщение такого фундаментального закона, как кирхгофовский закон излучения, на общий случай произвольного соотношения между размерами тел и длиной волны (§ 17).

Заметим кстати, что закон Кирхгофа, строго говоря, относится уже к *неравновесным* условиям. Ведь речь идет об излучении нагретого тела в окружающую более холодную среду, т. е. температура не одинакова во всем пространстве. Если теряемая телом энергия не компенсируется, то процесс даже нестационарен: тело «высвечивается» и остывает. Разумеется, нетрудно создать и стационарные условия, поддерживая температуру тела постоянной путем подвода энергии от какого-либо стороннего источника, как это происходит, скажем, в лампочке накаливания. Ясно, однако, что и такое стационарное состояние все равно термодинамически неравновесно: в теле происходит одностороннее преобразование энергии и протекают явления переноса (например, связанные с теплопроводностью). В этих условиях применение равновесных законов (в том числе ФДТ) уже не вполне строго, но оно оправдано, если, как часто бывает, роль явлений переноса еще невелика.

Последнее замечание, которое уместно здесь сделать, касается функций распределения вероятностей флуктуационного поля. Казалось бы, нахождение этих функций — более сложная задача, чем вычисление моментов, но во многих случаях дело обстоит как раз наоборот. Флуктуационное поле складывается из множества *микрополей*, порожденных флуктуациями заряда и тока в разных элементах объема. Если эти вклады можно считать статистически независимыми и «равноправными» по величине (например, по их доле в суммарной дисперсии), то в силу центральной предельной теоремы теории вероятностей суммарное поле будет *гауссовым*. Тем самым, для его полного статистического описания достаточно знать только функции корреляции его напряженностей (поскольку средние значения этих напряженностей равны нулю).

## § 14. Стохастические уравнения Максвелла

Исходные уравнения Максвелла

$$\operatorname{rot} \mathbf{H} = \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t} + \frac{4\pi}{c} \mathbf{j}, \quad \operatorname{rot} \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} \quad (14.1)$$

необходимо, как известно, дополнить так называемыми *материальными уравнениями*, связывающими индукции  $\mathbf{D}$  и  $\mathbf{B}$  с напряженностями  $\mathbf{E}$  и  $\mathbf{H}$ . Эти уравнения описывают электродинамические свойства среды и могут быть в общем случае нелинейны. Однако если задача о тепловых флуктуациях ставится *макроскопическими* уравнениями (а именно так она ставится во флуктуационной электродинамике) и если мы интересуемся только термодинамически *равновесными* флуктуациями, то макроскопическая нелинейность системы не играет роли, т. е. достаточно пользоваться *линеаризованными* уравнениями. Это обстоятельство было разъяснено ранее для дискретных систем (ч. I, § 54), но оно остается в силе и для распределенных систем. Другими словами, сохраняя полную общность, мы можем исходить из *линеаризованных* материальных уравнений. Последние можно было бы взять наиболее общими, учитывающими наличие неоднородности и анизотропии среды, временную и пространственную дисперсию, а также движение среды. Однако для упрощения выкладок и для более четкого выделения принципиальных моментов мы ограничимся случаем неподвижной и поначалу изотропной среды, обладающей только временной нелокальностью (частотной дисперсией). В этом частном случае среду можно описать материальными уравнениями вида (см. сноску на стр. 173)

$$\begin{aligned} \mathbf{D}(t, \mathbf{r}) &= \int_{-\infty}^{+\infty} \varepsilon(t-t', \mathbf{r}) \mathbf{E}(t', \mathbf{r}) dt', \\ \mathbf{B}(t, \mathbf{r}) &= \int_{-\infty}^{+\infty} \mu(t-t', \mathbf{r}) \mathbf{H}(t', \mathbf{r}) dt', \end{aligned} \quad (14.2)$$

где  $\varepsilon$  и  $\mu$  — диэлектрическая и магнитная проницаемости.

Переходя к стохастическим максвелловским уравнениям, введем теперь в правые части уравнений (14.1) случайные сторонние источники, которые «вызывают» тепловые флуктуации напряженностей поля и всех связанных с ними электродинамических величин. Эти распределенные источники играют здесь такую же вспомогательную роль эквивалентных ланжевеновских «сил», как интегральные э. д. с. Найквиста в электрических цепях с сосредоточенными параметрами (ч. I, § 54).

Флуктуационные источники можно выбирать различным образом, в частности в виде объемных плотностей электрических ( $\mathbf{j}$ )

и магнитных ( $j_m$ ) токов<sup>1)</sup>. Стохастические уравнения Максвелла примут тогда вид

$$\operatorname{rot} \mathbf{H} = \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t} + \frac{4\pi}{c} \mathbf{j} + \frac{4\pi}{c} \mathbf{j}_e, \quad \operatorname{rot} \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} - \frac{4\pi}{c} \mathbf{j}_m. \quad (14.3)$$

Можно было бы воспользоваться вместо токов сторонними индукциями  $\mathbf{D}$  и  $\mathbf{B}$ , положив

$$\mathbf{j}_e = \frac{1}{4\pi} \frac{\partial \mathbf{D}}{\partial t}, \quad \mathbf{j}_m = \frac{1}{4\pi} \frac{\partial \mathbf{B}}{\partial t}, \quad (14.4)$$

что, как мы увидим далее, оказывается в некоторых случаях целесообразным. Поля флуктуационных источников  $\mathbf{j}_e$  и  $\mathbf{j}_m$  (равно как и  $\mathbf{D}$ ,  $\mathbf{B}$ ) мы будем считать *стационарными* во времени.

Обратим внимание на вытекающее из (14.3) уравнение баланса энергии (теорему Пойнтинга), которое мы запишем для *полного* поля<sup>2)</sup>. Умножив первое уравнение (14.1) на  $\mathbf{E}$ , второе — на  $-\mathbf{H}$ ; сложив результаты и взяв интеграл по объему  $V$ , ограниченному поверхностью  $\Sigma$ , получаем

$$\int_V (\mathbf{E} \operatorname{rot} \mathbf{H} - \mathbf{H} \operatorname{rot} \mathbf{E}) d^3r = \frac{4\pi}{c} \left\{ \frac{dU}{dt} + Q + \int_V (\mathbf{E} \mathbf{j}_e + \mathbf{H} \mathbf{j}_m) d^3r \right\},$$

где

$$\frac{dU}{dt} = \frac{1}{4\pi} \int_V \left( \mathbf{E} \frac{\partial \mathbf{D}}{\partial t} + \mathbf{H} \frac{\partial \mathbf{B}}{\partial t} \right) d^3r, \quad Q = \int_V \mathbf{j} \mathbf{E} d^3r.$$

$dU/dt$  — скорость изменения заключенной в объеме  $V$  электромагнитной энергии, а  $Q$  — диссипируемая в этом объеме мощность. Интеграл в левой части равенства пропорционален потоку энергии (потоку вектора Пойнтинга) через поверхность  $\Sigma$ , так что для *полного* поля он обращается в нуль. Поэтому  $dU/dt \equiv Q$  представляет собой не что иное, как диссипируемую в объеме  $V$  мощность. После усреднения по ансамблю случайных «сил»  $\mathbf{j}_e$  и  $\mathbf{j}_m$ , если учесть, что для *полного* и *стационарного* поля  $d\langle U \rangle/dt = 0$ , мы получим для средней диссипируемой мощности выражение

$$\langle Q \rangle = - \int_V \langle \mathbf{E} \mathbf{j}_e + \mathbf{H} \mathbf{j}_m \rangle d^3r. \quad (14.5)$$

<sup>1)</sup> Использование магнитных токов является, конечно, лишь формальным приемом, но для многих задач оно удобно (см., например, [2]).

<sup>2)</sup> Напомним, что *полным* называется в электродинамике поле, равное нулю за пределами некоторой *замкнутой* поверхности  $\Sigma$ . Возможность построения такой поверхности обусловлена либо наличием идеально отражающих оболочек, либо достаточно быстрым убыванием поля на бесконечности (поверхность  $\Sigma$  может быть частично или целиком бесконечно удаленной).



Нас будут интересовать спектральные амплитуды стационарных во времени электродинамических величин, т. е. пространственные поля трансформант Фурье

$$E(\omega, r) = \frac{1}{2\pi} \int_{-\infty}^{\infty} E(t, r) e^{i\omega t} dt,$$

и аналогично для  $H$ ,  $D$ ,  $B$ ,  $j$ ,  $j_e$  и  $j_m$ . Для спектральных амплитуд уравнения (14.3) принимают вид

$$\operatorname{rot} H = -ikD + \frac{4\pi}{c} j_e, \quad \operatorname{rot} E = ikB - \frac{4\pi}{c} j_m, \quad (14.6)$$

где  $k = \omega/c$ , а материальные уравнения (14.2) — вид

$$\begin{aligned} D(\omega, r) &= \varepsilon(\omega, r) E(\omega, r), \\ B(\omega, r) &= \mu(\omega, r) H(\omega, r), \end{aligned} \quad (14.7)$$

где

$$\left. \begin{aligned} \varepsilon(\omega, r) \\ \mu(\omega, r) \end{aligned} \right\} = \int_0^{\infty} \left. \begin{aligned} \varepsilon(\tau, r) \\ \mu(\tau, r) \end{aligned} \right\} e^{i\omega\tau} d\tau. \quad (14.8)$$

Заметим, что член  $4\pi j/c$  мы объединили в (14.6) с членом  $ikD$ . Это означает включение электрической проводимости в  $\varepsilon$  или, точнее, в ее мнимую часть  $\varepsilon''$  ( $\varepsilon = \varepsilon' + i\varepsilon''$ ). Вводя  $\mu$  ( $\mu = \mu' + i\mu''$ ) можно учесть и магнитные потери.

Система неоднородных линейных уравнений (14.6), (14.7) однозначно определяет спектральные амплитуды напряженностей и индукций флуктуационного поля при заданных граничных условиях и (если это нужно) условиях на бесконечности, а также, разумеется, при заданных случайных источниках  $j_e(\omega, r)$  и  $j_m(\omega, r)$ . Таким образом, это задача со случайными источниками (схема I по классификации, введенной в § 8), вероятностные свойства которых должны быть известны. Так как уравнения линейны, для нахождения вторых ( $n$ -х) моментов флуктуационного поля надо знать только вторые же ( $n$ -е) моменты источников. Эти вторые моменты, если речь идет о флуктуациях около состояния термодинамического равновесия, полностью определены (как и в случае сосредоточенных систем) флуктуационно-диссипационной теоремой (ФДТ), которую нетрудно распространить и на *распределенные* системы (поля). Поэтому мы теперь оставим на время уравнения Максвелла и обратимся к указанному обобщению ФДТ.

### § 15. Равновесные тепловые флуктуации в непрерывных диссипативных системах

В ч. 1, § 54 была приведена формулировка ФДТ для *дискретных* систем, состояние которых описывается совокупностью какого-то конечного числа  $n$  обобщенных координат  $q_i(t)$ . При помощи этой дискретной формы ФДТ можно получать результаты, относящиеся и к распределенным системам, — либо косвенным путем, как сделано в задаче 1, либо используя искусственную дискретизацию непрерывной системы, например разбиение пространства, занимаемого полем, на малые ячейки и замену дифференциальных операторов разностными. Именно с помощью последнего способа дискретная ФДТ применена к стохастическим уравнениям Максвелла в книге [3], гл. 13. Однако представляет больший интерес получить *общую* форму ФДТ для *непрерывных* систем, состояние которых описывается случайными полями — функциями времени и точки пространства. Такое обобщение, удобное для приложений теоремы к полям равновесных тепловых флуктуаций любой физической природы — как электромагнитным, так и механическим (в аэро- и гидродинамике, теории упругости), температурным и энтропийным и т. п., — можно осуществить следующим образом [4].

Пусть флуктуации в распределенной системе, занимающей объем  $V$ , описываются стационарным и однородным полем  $\xi(t, \mathbf{r})$ . Для простоты мы берем случай одномерного поля<sup>1)</sup>. Обобщение на многомерное поле  $\xi^{(i)}(t, \mathbf{r})$  ( $i = 1, 2, \dots$ ) легко может быть получено, и соответствующий результат будет далее приведен. Пусть  $f(t, \mathbf{r})$  — объемная плотность поля обобщенной силы, сопряженной в лагранжевом смысле с обобщенной координатой  $\xi(t, \mathbf{r})$ , причем спектральные амплитуды  $\xi(\omega, \mathbf{r})$  и  $f(\omega, \mathbf{r})$  связаны между собой посредством линейных пространственных операторов  $\hat{A}$  и  $\hat{A}^{-1}$ :

$$\xi(\omega, \mathbf{r}) = \hat{A}f(\omega, \mathbf{r}), \quad f(\omega, \mathbf{r}) = \hat{A}^{-1}\xi(\omega, \mathbf{r}). \quad (15.1)$$

$\hat{A}$  и  $\hat{A}^{-1}$  — взаимно обратные операторы, т. е. их произведение равно единичному оператору  $\hat{E}$  ( $\hat{A}\hat{A}^{-1} = \hat{A}^{-1}\hat{A} = \hat{E}$ ), который оставляет функцию неизменной ( $\hat{E}f(\mathbf{r}) = f(\mathbf{r})$ ). В приложениях обычно задан оператор  $\hat{A}^{-1}$ , являющийся большей частью дифференциальным.

Если, например, мы имеем дело с волновым уравнением

$$\Delta \xi - \frac{1}{c^2} \frac{\partial^2 \xi}{\partial t^2} = -4\pi f(t, \mathbf{r}),$$

<sup>1)</sup> Одномерного в смысле числа описывающих его функций  $\xi$ .

частное решение которого равно, как известно,

$$\xi(t, \mathbf{r}) = \int f\left(t - \frac{R}{c}, \mathbf{r}'\right) \frac{d^3 r'}{R} \quad (R = |\mathbf{r} - \mathbf{r}'|),$$

то для спектральных амплитуд получаем ( $k = \omega/c$ )

$$(\Delta + k^2) \xi(\omega, \mathbf{r}) = -4\pi f(\omega, \mathbf{r}),$$

$$\xi(\omega, \mathbf{r}) = \int f(\omega, \mathbf{r}') \frac{e^{ikR}}{R} d^3 r'.$$

Таким образом,

$$\hat{A}^{-1} = -\frac{\Delta + k^2}{4\pi}, \quad \hat{A}\{\dots\} = \int \{\dots\} \frac{e^{ikR}}{R} d^3 r'.$$

Допустим теперь, что мы располагаем полной ортонормированной системой вещественных функций  $\varphi_j(\mathbf{r})$ , определенной для области  $V$ :

$$\int_V \varphi_j(\mathbf{r}) \varphi_k(\mathbf{r}) d^3 r = \delta_{jk}. \quad (15.2)$$

Разложим  $\xi(t, \mathbf{r})$  и  $f(t, \mathbf{r})$  по функциям  $\varphi_j(\mathbf{r})$ :

$$\xi(t, \mathbf{r}) = \sum_j \xi_j(t) \varphi_j(\mathbf{r}), \quad f(t, \mathbf{r}) = \sum_j f_j(t) \varphi_j(\mathbf{r}), \quad (15.3a, б)$$

или в спектральной форме:

$$\xi(\omega, \mathbf{r}) = \sum_j \xi_j(\omega) \varphi_j(\mathbf{r}), \quad f(\omega, \mathbf{r}) = \sum_j f_j(\omega) \varphi_j(\mathbf{r}), \quad (15.4a, б)$$

где

$$\xi_j(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \xi_j(t) e^{i\omega t} dt,$$

и аналогично для  $f_j(\omega)$ . Заметим, что согласно (15.2) и (15.3) средняя мощность, диссипируемая в объеме  $V$  под действием силы  $f(t, \mathbf{r})$ , записывается в виде

$$\langle Q \rangle = \int \left\langle f \frac{\partial \xi}{\partial t} \right\rangle d^3 r = - \int \left\langle \xi \frac{\partial f}{\partial t} \right\rangle d^3 r = - \sum_j \left\langle \xi_j(t) \frac{\partial f_j(t)}{\partial t} \right\rangle. \quad (15.5)$$

Принимая  $\xi_j(t)$  за те дискретные переменные, которые описывают флуктуации в рассматриваемой системе, естественно распространить дискретную ФДТ на спектральные амплитуды  $\xi_j(\omega)$  и  $f_j(\omega)$ . Единственное допущение, которое необходимо при этом сделать, заключается в том, что ФДТ остается в силе для бесконечного множества дискретных переменных  $\xi_j(t)$ . Опираясь на это допущение, мы и найдем теперь пространственные функции корреляции для спектральных амплитуд  $\xi(\omega, \mathbf{r})$  и  $f(\omega, \mathbf{r})$ .

Как связаны между собой спектральные амплитуды  $\xi_j(\omega)$  и  $f_j(\omega)$ ? Подставив (15.3а, б) в первое уравнение (15.1), получаем

$$\sum_k \xi_k(\omega) \varphi_k(r) = \sum_k f_k(\omega) \hat{A} \varphi_k(r).$$

Умножим это равенство слева на  $\varphi_j(r)$  и проинтегрируем по объему  $V$ . Учитывая (15.2), находим

$$\xi_j(\omega) = \sum_k f_k(\omega) \int_V \varphi_j(r) \hat{A} \varphi_k(r) d^3r \equiv \sum_k \alpha_{jk}(\omega) f_k(\omega), \quad (15.6a)$$

где величины

$$\alpha_{jk}(\omega) = \int_V \varphi_j(r) \hat{A} \varphi_k(r) d^3r \quad (15.7a)$$

представляют собой, очевидно, не что иное, как коэффициенты разложения функций  $\hat{A} \varphi_k(r)$  по базисным функциям  $\varphi_j(r)$ , т. е.

$$\hat{A} \varphi_k(r) = \sum_j \alpha_{jk} \varphi_j(r). \quad (15.8a)$$

Точно такие же операции над вторым уравнением (15.1) приводят к обратной (15.6а) системе уравнений:

$$f_j(\omega) = \sum_k \alpha_{jk}^{-1}(\omega) \xi_k(\omega), \quad (15.6б)$$

где

$$\alpha_{jk}^{-1}(\omega) = \int_V \varphi_j(r) \hat{A}^{-1} \varphi_k(r) d^3r, \quad (15.7б)$$

т. е.  $\alpha_{jk}^{-1}$  — коэффициенты разложения по  $\varphi_j(r)$  функции  $\hat{A}^{-1} \varphi_k(r)$ :

$$\hat{A}^{-1} \varphi_k(r) = \sum_j \alpha_{jk}^{-1} \varphi_j(r). \quad (15.8б)$$

Величины (15.7а, б), являющиеся коэффициентами в уравнениях (15.6а, б), связывающих  $\xi_j(\omega)$  и  $f_j(\omega)$ , образуют, таким образом, элементы бесконечномерной матрицы *обобщенной восприимчивости*  $\{\alpha_{jk}^{-1}\}$  и обратной ей матрицы  $\{\alpha_{jk}\}$ .

Дискретная ФДТ дает следующие выражения для спектральных плотностей переменных  $\xi_j(t)$  и ланжевеновских сил  $f_j(t)$  (ч. I, (54.18) и (54.22); отметим, что перед правыми частями теперь стоит знак минус, поскольку мы пользовались ранее спектральным разложением с  $e^{i\omega t}$ , а в данной книге перешли к разложению по  $e^{-i\omega t}$ ):

$$\langle \xi_j(\omega) \xi_k^*(\omega) \rangle = - \frac{i\theta(\omega, T)}{2\pi\omega} (\alpha_{jk} - \alpha_{kj}^*), \quad (15.9a)$$

$$\langle f_j(\omega) f_k^*(\omega) \rangle = - \frac{i\theta(\omega, T)}{2\pi\omega} (\alpha_{kj}^{-1*} - \alpha_{jk}^{-1}), \quad (15.9б)$$

где  $\Theta(\omega, T)$  — средняя энергия квантового осциллятора:

$$\Theta(\omega, T) = \hbar\omega \left\{ \frac{1}{e^{\hbar\omega/kT} - 1} + \frac{1}{2} \right\} = \frac{\hbar\omega}{2} \coth \frac{\hbar\omega}{2kT}. \quad (15.10)$$

Формулами (15.9а, б) мы теперь и воспользуемся.

Согласно (15.4а) и (15.9а)

$$\begin{aligned} \langle \xi(\omega, \mathbf{r}) \xi^*(\omega, \mathbf{r}') \rangle &= \sum_{j, k} \langle \xi_j(\omega) \xi_k^*(\omega) \rangle \varphi_j(\mathbf{r}) \varphi_k(\mathbf{r}') = \\ &= -\frac{i\Theta(\omega, T)}{2\pi\omega} \sum_{j, k} (\alpha_{jk} - \alpha_{kj}^*) \varphi_j(\mathbf{r}) \varphi_k(\mathbf{r}'). \end{aligned}$$

При помощи (15.8а) можно записать это выражение в виде

$$-\langle \xi(\omega, \mathbf{r}) \xi^*(\omega, \mathbf{r}') \rangle = -\frac{i\Theta(\omega, T)}{2\pi\omega} \sum_j \{ \varphi_j(\mathbf{r}') \hat{A} \varphi_j(\mathbf{r}) - \varphi_j(\mathbf{r}) \hat{A}^* \varphi_j(\mathbf{r}') \},$$

причем оператор  $\hat{A}$  действует на функции точки  $\mathbf{r}$ , а  $\hat{A}^*$  — на функции точки  $\mathbf{r}'$ . Поэтому оба они могут быть вынесены за знак суммы, которая оказывается тогда разложением дельта-функции от  $\mathbf{r} - \mathbf{r}'$  по базисным функциям:

$$\sum_j \varphi_j(\mathbf{r}) \varphi_j(\mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}').$$

В результате

$$\langle \xi(\omega, \mathbf{r}) \xi^*(\omega, \mathbf{r}') \rangle = -\frac{i\Theta(\omega, T)}{2\pi\omega} (\hat{A} - \hat{A}^*) \delta(\mathbf{r} - \mathbf{r}'). \quad (15.11а)$$

Совершенно аналогичным путем при помощи (15.4б), (15.9б) и (15.8б) получается пространственная функция корреляции для спектральной амплитуды силы  $f(\omega, \mathbf{r})$ :

$$\langle f(\omega, \mathbf{r}) f^*(\omega, \mathbf{r}') \rangle = -\frac{i\Theta(\omega, T)}{2\pi\omega} (\hat{A}^{-1*} - \hat{A}^{-1}) \delta(\mathbf{r} - \mathbf{r}'). \quad (15.11б)$$

Поскольку дельта-функция зависит лишь от разности  $\mathbf{r} = \mathbf{r} - \mathbf{r}'$ , оба оператора  $\hat{A}$  и  $\hat{A}^*$  можно выразить через компоненты  $\mathbf{p}$ . Таким образом можно и дифференциальные операторы  $\hat{A}^{-1}$  и  $\hat{A}^{-1*}$  (первый содержит  $\nabla$ , а второй  $\nabla'$ ) записать через оператор  $\nabla_p = \nabla = -\nabla'$ .

Формулы (15.11а, б), в которых уже никак не проявляются форма и размеры области  $V$ , и представляют ФДТ для одномерного однородного поля  $\xi(t, \mathbf{r})$  и соответствующего ему поля ланжевеновской силы  $f(t, \mathbf{r})$ .

Обобщение полученного результата на многомерное поле, т. е. на случай системы однородных и однородно связанных между собой полей  $\xi^{(i)}(t, \mathbf{r})$ , почти очевидно. Вместо уравнений (15.1)

теперь будут *системы* линейных уравнений

$$\begin{aligned}\xi^{(j)}(\omega, \mathbf{r}) &= \sum_k \hat{A}_{jk} f^{(k)}(\omega, \mathbf{r}), \\ f^{(j)}(\omega, \mathbf{r}) &= \sum_k \hat{A}_{jk}^{-1} \xi^{(k)}(\omega, \mathbf{r}),\end{aligned}\quad (15.12)$$

где  $\hat{A}_{jk}$  и  $\hat{A}_{jk}^{-1}$  — элементы прямой и обратной *операторных матриц*, а вместо выражения (15.5) для мощности, развиваемой силами  $f^{(j)}(t, \mathbf{r})$ , будет сумма

$$\begin{aligned}\langle Q \rangle &= \sum_j \int_V \left\langle f^{(j)}(t, \mathbf{r}) \frac{\partial \xi^{(j)}(t, \mathbf{r})}{\partial t} \right\rangle d^3r = \\ &= - \sum_j \int_V \left\langle \xi^{(j)}(t, \mathbf{r}) \frac{\partial f^{(j)}(t, \mathbf{r})}{\partial t} \right\rangle d^3r.\end{aligned}\quad (15.13)$$

Наконец, окончательные формулы (15.11а, б) заменяются на корреляционные *матрицы* спектральных амплитуд  $\xi^{(j)}(\omega, \mathbf{r})$  и  $f^{(j)}(\omega, \mathbf{r})$ :

$$\langle \xi^{(j)}(\omega, \mathbf{r}) \xi^{(k)*}(\omega, \mathbf{r}') \rangle = - \frac{i\theta(\omega, T)}{2\pi\omega} (\hat{A}_{jk} - \hat{A}_{kj}^*) \delta(\mathbf{r}), \quad (15.14a)$$

$$\langle f^{(j)}(\omega, \mathbf{r}) f^{(k)*}(\omega, \mathbf{r}') \rangle = - \frac{i\theta(\omega, T)}{2\pi\omega} (\hat{A}_{kj}^{-1*} - \hat{A}_{jk}^{-1}) \delta(\mathbf{r}), \quad (15.14б)$$

где по-прежнему  $\mathbf{r} = \mathbf{r} - \mathbf{r}'$ .

Итак, в распределенных системах, как и в дискретных, ФДТ выражает корреляционные функции (матрицы)  $\omega$ -амплитуд<sup>1)</sup> совершенно регулярным образом. Эти функции однозначно определены видом операторов  $\hat{A}_{jk}$  и  $\hat{A}_{jk}^{-1}$ , т. е. *самими линеаризованными макроскопическими уравнениями* рассматриваемой системы. Если это уравнения Максвелла, то формулы (15.14) позволяют найти пространственную корреляцию напряженностей флуктуационного поля и тех сторонних полей (индукций или токов), которые «создают» флуктуационное электромагнитное поле. Если же мы имеем дело, например, с уравнениями теории упругости, то при помощи тех же формул (15.14) определяется корреляция тепловых флуктуаций деформаций и температуры, равно как и соответствующих сторонних напряжений и источников тепла [5] и т. д. Более точно: корреляционные матрицы флуктуационных и сторонних тепловых полей определяются *антиэрмитовыми частями* соответствующих линейных операторов, или иначе говоря, обусловлены *диссипативными* свойствами системы. В отсутствие поглощения энергии, т. е. в отсутствие диссипа-

<sup>1)</sup> А тем самым и  $\omega$ -амплитуд, если  $\omega$ -амплитуды разложить в пространственные интегралы Фурье (см. задачу 6).

ции, операторы или операторные матрицы будут эрмитовыми и разности, стоящие в правых частях любой из формул, выражающих ФДТ, обращаются в нуль. Это означает, что в среде (или участке среды), не обладающей потерями, сторонних источников флуктуаций нет и, следовательно, нет и вклада от такого участка среды во флуктуационное поле. В формулах (15.11а) и (15.14а) учтем *только этот вклад*, и он в данном случае тоже обращается в нуль. Но флуктуационное поле может создаваться источниками, локализованными *вне* данного непоглощающего объема среды, т. е. может приходить в этот объем извне. Поэтому корреляционные функции *флуктуационного поля*, вообще говоря, отличны от нуля и в тех областях пространства, где среда не обладает поглощением.

### § 16. Корреляция сторонних тепловых источников в электродинамике

Вернемся к спектральным уравнениям Максвелла (14.6), исключив из них при помощи материальных уравнений (14.7) индукции  $\mathbf{D}$  и  $\mathbf{B}$ :

$$\operatorname{rot} \mathbf{H} = -ik\epsilon\mathbf{E} + \frac{4\pi}{c} \mathbf{j}_e, \quad \operatorname{rot} \mathbf{E} = ik\mu\mathbf{H} - \frac{4\pi}{c} \mathbf{j}_m. \quad (16.1)$$

Состояние рассматриваемой системы — флуктуационного электромагнитного поля — описывается 6-вектором напряженностей  $\xi^{(j)} = \{\mathbf{E}, \mathbf{H}\}$ . Каков при этом 6-вектор сопряженных с  $\xi^{(j)}$  по Лагранжу сторонних сил  $f^{(j)}$ ?

Согласно общему выражению (15.13) средняя диссипируемая мощность  $\langle Q \rangle$  равна (со знаком минус) сумме произведений координат на *скорости* изменения сил. Сопоставляя с (15.13) формулу (14.5), нетрудно усмотреть, что *сторонние токи*  $\mathbf{j}_e$  и  $\mathbf{j}_m$  не являются сопряженными «силами» для «координат»  $\xi^{(j)} = \{\mathbf{E}, \mathbf{H}\}$ . Если же воспользоваться *сторонними индукциями*  $\mathbf{D}$  и  $\mathbf{B}$  и подставить в (14.5) выражения (14.4), то мы получим для  $\langle Q \rangle$  формулу как раз вида (15.13):

$$\langle Q \rangle = -\frac{1}{4\pi} \int_V \left\langle \mathbf{E} \frac{\partial \mathbf{D}}{\partial t} + \mathbf{H} \frac{\partial \mathbf{B}}{\partial t} \right\rangle d^3r.$$

Таким образом, для напряженностей  $\xi^{(j)} = \{\mathbf{E}, \mathbf{H}\}$  «силами» являются величины  $f^{(j)} = \{\mathbf{D}/4\pi, \mathbf{B}/4\pi\}$  и общие формулы (15.14а, б), выражающие ФДТ, следует применять именно к таким  $\xi^{(j)}$  и  $f^{(j)}$ .

Разумеется, получив корреляционную матрицу для спектральных амплитуд сторонних индукций  $\mathbf{D}$  и  $\mathbf{B}$ , мы тотчас же сможем написать такую матрицу и для спектральных амплитуд

сторонних токов, поскольку, в силу (14.4),

$$j_e = -\frac{i\omega}{4\pi} \mathfrak{D}, \quad j_m = -\frac{i\omega}{4\pi} \mathfrak{B}. \quad (16.2)$$

Подставив (16.2) в (16.1), запишем уравнения (16.1) в компонентах:

$$\begin{aligned} \frac{\mathfrak{D}_1}{4\pi} &= -\frac{\varepsilon E_1}{4\pi} + \frac{1}{4\pi i k} (\nabla_2 H_3 - \nabla_3 H_2) \mathcal{O}, \\ \frac{\mathfrak{B}_1}{4\pi} &= -\frac{\mu H_1}{4\pi} - \frac{1}{4\pi i k} (\nabla_2 E_3 - \nabla_3 E_2) \mathcal{O}, \end{aligned}$$

где  $k = \omega/c$ , а знак  $\mathcal{O}$  обозначает круговую перестановку индексов 1, 2, 3. Сопоставление этих шести уравнений со второй группой уравнений (15.12), где, как мы помним, надо положить

$$\xi^{(j)} = \{E, H\}, \quad f^{(j)} = \{\mathcal{O}/4\pi, \mathfrak{B}/4\pi\} \quad (j = 1, \dots, 6),$$

дает нам матрицу операторов  $\hat{A}_{jk}^{-1}(\nabla)$ :

$$\hat{A}_{jk}^{-1}(\nabla) = \left\{ \begin{array}{c|c} -\frac{\varepsilon}{4\pi} \delta_{\alpha\beta} & \frac{1}{4\pi i k} \begin{pmatrix} 0 & \nabla_3 & -\nabla_2 \\ -\nabla_3 & 0 & \nabla_1 \\ \nabla_2 & -\nabla_1 & 0 \end{pmatrix} \\ \hline -\frac{1}{4\pi i k} \begin{pmatrix} 0 & \nabla_3 & -\nabla_2 \\ -\nabla_3 & 0 & \nabla_1 \\ \nabla_2 & -\nabla_1 & 0 \end{pmatrix} & -\frac{\mu}{4\pi} \delta_{\alpha\beta} \end{array} \right\} \\ (\alpha, \beta = 1, 2, 3).$$

Транспонированная комплексно сопряженная матрица, зависящая от  $\nabla'$ , будет

$$\hat{A}_{kj}^{-1*}(\nabla') = \left\{ \begin{array}{c|c} -\frac{\varepsilon^*}{4\pi} \delta_{\alpha\beta} & -\frac{1}{4\pi i k} \begin{pmatrix} 0 & \nabla'_3 & -\nabla'_2 \\ -\nabla'_3 & 0 & \nabla'_1 \\ \nabla'_2 & -\nabla'_1 & 0 \end{pmatrix} \\ \hline \frac{1}{4\pi i k} \begin{pmatrix} 0 & \nabla'_3 & -\nabla'_2 \\ -\nabla'_3 & 0 & \nabla'_1 \\ \nabla'_2 & -\nabla'_1 & 0 \end{pmatrix} & -\frac{\mu^*}{4\pi} \delta_{\alpha\beta} \end{array} \right\} \\ (\alpha, \beta = 1, 2, 3).$$

Учитывая, что  $\nabla' = -\nabla$  (поскольку операторы действуют на функцию от  $\mathbf{r} - \mathbf{r}'$ ), получаем

$$\hat{A}_{kj}^{-1*}(\nabla') - \hat{A}_{jk}^{-1}(\nabla) = \left\{ \begin{array}{c|c} \frac{\varepsilon - \varepsilon^*}{4\pi} \delta_{\alpha\beta} & 0 \\ \hline 0 & \frac{\mu - \mu^*}{4\pi} \delta_{\alpha\beta} \end{array} \right\} \quad (\alpha, \beta = 1, 2, 3).$$



Наконец, подстановка этого результата в формулу (15.146), выражающую ФДТ для ланжевеновских сил, даст нам корреляционные функции сторонних индукций  $\mathfrak{D}$  и  $\mathfrak{B}$ :

$$\begin{aligned}\langle \mathfrak{D}_\alpha(\omega, \mathbf{r}) \mathfrak{D}_\beta^*(\omega, \mathbf{r}') \rangle &= -\frac{2i\Theta(\omega, T)}{\omega} (\epsilon - \epsilon^*) \delta_{\alpha\beta} \delta(\mathbf{r} - \mathbf{r}'), \\ \langle \mathfrak{B}_\alpha(\omega, \mathbf{r}) \mathfrak{B}_\beta^*(\omega, \mathbf{r}') \rangle &= -\frac{2i\Theta(\omega, T)}{\omega} (\mu - \mu^*) \delta_{\alpha\beta} \delta(\mathbf{r} - \mathbf{r}'), \\ \langle \mathfrak{D}_\alpha(\omega, \mathbf{r}) \mathfrak{B}_\beta^*(\omega, \mathbf{r}') \rangle &= 0.\end{aligned}\quad (16.3)$$

В более общем случае *анизотропной* или гиротропной среды, описываемой *тензорами* проницаемостей  $\epsilon_{\alpha\beta}$  и  $\mu_{\alpha\beta}$ , мы получили бы те же формулы (16.3), но с заменой диагональных тензоров  $\epsilon\delta_{\alpha\beta}$  и  $\mu\delta_{\alpha\beta}$  на  $\epsilon_{\alpha\beta}$  и  $\mu_{\alpha\beta}$ , а  $\epsilon^*\delta_{\alpha\beta}$  и  $\mu^*\delta_{\alpha\beta}$  на  $\epsilon_{\beta\alpha}^*$  и  $\mu_{\beta\alpha}^*$ .

Пользуясь (16.2) и (16.3), нетрудно написать корреляционные функции сторонних токов. Мы приведем эти функции сразу для случая анизотропной среды <sup>1</sup>:

$$\begin{aligned}\langle j_{\alpha\alpha}(\omega, \mathbf{r}) j_{\beta\beta}^*(\omega, \mathbf{r}') \rangle &= -\frac{i\omega\Theta(\omega, T)}{8\pi^2} (\epsilon_{\alpha\beta} - \epsilon_{\beta\alpha}^*) \delta(\mathbf{r} - \mathbf{r}'), \\ \langle j_{m\alpha}(\omega, \mathbf{r}) j_{m\beta}^*(\omega, \mathbf{r}') \rangle &= -\frac{i\omega\Theta(\omega, T)}{8\pi^2} (\mu_{\alpha\beta} - \mu_{\beta\alpha}^*) \delta(\mathbf{r} - \mathbf{r}'), \\ \langle j_{\alpha\alpha}(\omega, \mathbf{r}) j_{m\beta}^*(\omega, \mathbf{r}') \rangle &= 0.\end{aligned}\quad (16.4)$$

Менее формальный вывод этих формул дан в задаче 2.

Итак, электрические и магнитные источники флуктуационного поля пространственно не коррелированы между собой, а радиус пространственной корреляции тех и других порознь равен нулю (дельта-функция  $\delta(\mathbf{r} - \mathbf{r}')$ ). Для сред без пространственной дисперсии, которыми мы здесь и ограничиваемся, этот последний результат представляется очевидным. Действительно, фактический радиус корреляции *источников* теплового поля в таких средах может быть только микровеличиной (порядка, например, межатомных расстояний), и поэтому в макроскопической теории, рассматривающей среду как сплошную, он и должен быть равен нулю. Напротив, в средах с *пространственной дисперсией* материальные уравнения нелокальны не только по  $t$ , но и по  $\mathbf{r}$ , т. е. вместо (14.2) будут уравнения, содержащие операторы как по  $t$ , так и по  $\mathbf{r}$ <sup>2</sup>). Это приводит к

<sup>1</sup> Формулы (16.4) были получены Л. Д. Ландау и Е. М. Лифшицем [3] при помощи ФДТ в ее дискретной форме. На основе непрерывной формы ФДТ формулы вида (16.4) были выведены в [4]. В дальнейшем, независимо от полученной в [4] общей непрерывной формы ФДТ, формулы вида (16.4) выводили многие авторы (см., например, ссылки [23, 55—62] в книге [6]).

<sup>2</sup> Соответственно уравнения (14.7) тоже не будут алгебраическими, а будут содержать пространственные операторы. Алгебраическими будут лишь уравнения для  $\omega$ -амплитуд.

отличному от нуля радиусу корреляции источников флуктуационного поля, имеющему тот же порядок величины, что и размер области нелокальности в материальных уравнениях (см. [6], § 4).

Для кусочно-однородных сред, т. е. при наличии резких границ раздела между средами, формулы (16.3) и (16.4) справедливы (из-за дельта-корреляции) вплоть до самых границ раздела. При наличии же пространственной нелокальности положение меняется: вблизи от границ раздела, а именно в слое, толщина которого порядка радиуса корреляции, поля сторонних источников неоднородны (их корреляционные функции зависят от  $\mathbf{r}$  и  $\mathbf{r}'$  в отдельности, а не от разности  $\mathbf{r} - \mathbf{r}'$ ) и анизотропны даже в том случае, когда сама среда изотропна.

Мы смогли, используя теорему (15.14б), сразу написать корреляционные функции источников флуктуационного поля только потому, что входящие в (15.14б) операторы  $\hat{A}_{\alpha\beta}^{-1}$  нам известны из самих уравнений Максвелла. Иначе обстоит дело с корреляцией «координат», т. е. напряженностей флуктуационного электромагнитного поля  $\mathbf{E}$  и  $\mathbf{H}$ . Для того чтобы написать корреляционные функции  $\mathbf{E}$  и  $\mathbf{H}$ , надо, согласно (15.14а), знать обратные операторы  $\hat{A}_{\alpha\beta}$ , т. е. надо «обернуть» уравнения Максвелла, выразив  $\mathbf{E}$  и  $\mathbf{H}$  через  $\mathbf{j}_e$  и  $\mathbf{j}_m$  (или  $\mathfrak{D}$  и  $\mathfrak{B}$ ). Другими словами, надо решить неоднородные уравнения Максвелла, считая распределение источников  $\mathbf{j}_e$  и  $\mathbf{j}_m$  произвольным, но заданным и налагая граничные условия, отвечающие данной конкретной задаче.

Если краевая задача решена, то тем самым мы получаем операторы  $\hat{A}_{\alpha\beta}$  (обычно в виде интегралов, распространенных на области пространства, в которых сторонние токи  $\mathbf{j}_e$  и  $\mathbf{j}_m$  отличны от нуля) и можем воспользоваться тогда «готовой» теоремой (15.14а). Можно поступить и иначе: составить из полученных решений для  $\mathbf{E}$  и  $\mathbf{H}$  интересующие нас билинейные комбинации и усреднить их. Под кратные интегралы, выражающие вторые моменты напряженностей, войдут при этом корреляционные функции сторонних токов  $\mathbf{j}_e$  и  $\mathbf{j}_m$ , вместо которых надо будет подставить уже полученные выражения (16.4). Оба способа вполне правомерны, и ряд флуктуационных электродинамических задач был решен именно таким путем [1, 7, 8].

Однако в обоих указанных вариантах это трудоемкий и неэкономный путь, так как для каждой конкретной задачи его надо проделывать заново. Попытаемся поэтому выяснить, к чему приводит ФДТ, если применить ее к решению уравнений Максвелла, записанному в общей интегральной форме — через функции Грина.

## § 17. Обобщенный закон Кирхгофа

Функции Грина представляют собой решения краевой задачи с элементарными *точечными* источниками поля. Введем их в нашем электродинамическом случае следующим образом.

Пусть  $j_{0e}$  — детерминированный электрический ток, сосредоточенный в некоторой точке  $A$  ( $\mathbf{r} = \mathbf{r}_A$ ) и направленный по постоянному единичному вектору  $\mathbf{l}$ :

$$j_{0e} = l \delta(\mathbf{r} - \mathbf{r}_A). \quad (17.1)$$

Можно считать ток  $j_{0e}$  поляризационным, т. е. рассматривать источник как *точечный электрический диполь* с моментом  $\mathbf{p} = -l/i\omega \mathbf{l}$ . Решение уравнений Максвелла, т. е. напряженности электромагнитного поля, создаваемого в данной системе тел и сред источником (17.1), мы обозначим через  $\mathbf{E}_{0e}(\mathbf{r}, \mathbf{r}_A, l)$ ,  $\mathbf{H}_{0e}(\mathbf{r}, \mathbf{r}_A, l)$  и для краткости будем называть это *детерминированное поле дифракционным*. Напряженности  $\mathbf{E}_{0e}$  и  $\mathbf{H}_{0e}$  и являются функциями Грина.

Наша цель состоит в том, чтобы выразить напряженности  $\mathbf{E}$  и  $\mathbf{H}$  флуктуационного поля, порождаемого случайными распределенными токами  $\mathbf{j}_e$  и  $\mathbf{j}_m$ , через эти токи и через функции Грина  $\mathbf{E}_{0e}$  и  $\mathbf{H}_{0e}$ . Проще всего сделать это при помощи электродинамической *теоремы взаимности*, которая связывает «накрест» напряженности и источники двух различных полей 1 и 2 в одной и той же системе тел и сред.

Пусть имеются два распределения сторонних электрических и магнитных токов  $\mathbf{j}_{1e}$ ,  $\mathbf{j}_{1m}$  и  $\mathbf{j}_{2e}$ ,  $\mathbf{j}_{2m}$ , которые создают при прочих одинаковых условиях поля  $\mathbf{E}_1$ ,  $\mathbf{H}_1$  и  $\mathbf{E}_2$ ,  $\mathbf{H}_2$  соответственно. Теорема взаимности выводится весьма просто (при очень широких допущениях) из двух систем уравнений Максвелла — одной для поля и источников 1 и другой для поля и источников 2 (см., например, [6]). Она гласит, что

$$\int_V (\mathbf{E}_1 \mathbf{j}_{2e} - \mathbf{H}_1 \mathbf{j}_{2m}) d^3r = \int_V (\mathbf{E}_2 \mathbf{j}_{1e} - \mathbf{H}_2 \mathbf{j}_{1m}) d^3r, \quad (17.2)$$

где  $V$  — объем полного поля. Фактически интегралы в (17.2) распространяются только на те области пространства, в которых сторонние токи отличны от нуля.

<sup>1</sup> Плотность поляризационного тока равна  $\partial \mathbf{P} / \partial t$ , где  $\mathbf{P}$  — объемная плотность поляризации. Для точечного диполя с моментом  $\mathbf{p}$  имеем  $\mathbf{P} = p \delta(\mathbf{r} - \mathbf{r}_A)$ , так что  $j_{0e} = \frac{\partial p}{\partial t} \delta(\mathbf{r} - \mathbf{r}_A)$ , а для спектральных амплитуд  $j_{0e} = -i\omega p \delta(\mathbf{r} - \mathbf{r}_A)$ . При  $p = -l/i\omega$  мы получаем (17.1).

Отождествим теперь в теореме (17.2) интересующее нас флуктуационное поле с полем 1:

$$j_{e1} = j_e, \quad j_{m1} = j_m, \quad E_1 = E, \quad H_1 = H, \quad (17.3)$$

а вспомогательное дифракционное поле — с полем 2:

$$j_{e2} = l \delta(r - r_A), \quad j_{m2} = 0, \quad E_2 = E_{oe}, \quad H_2 = H_{oe};$$

формула (17.2) дает тогда

$$\begin{aligned} \int_V E(r) l \delta(r - r_A) d^3r &= E_l(r_A) = \\ &= \int_V \{E_{oe}(r, r_A, l) j_e(r) - H_{oe}(r, r_A, l) j_m(r)\} d^3r, \end{aligned} \quad (17.4)$$

т. е. именно то, что нам нужно: компонента по направлению  $l$  электрической напряженности флуктуационного поля в точке  $A$  выражена через распределение случайных токов и через функции Грина — напряженности детерминированного поля, создаваемого электрическим диполем, направленным по  $l$  и находящимся в точке  $A$ .

Для того чтобы получить аналогичное выражение для компоненты магнитного флуктуационного поля, надо, сохраняя (17.3), воспользоваться в теореме взаимности (17.2) другим полем 2, а именно дифракционным полем  $E_{om}$ ,  $H_{om}$ , которое создается магнитным точечным током

$$j_{om} = l \delta(r - r_A)$$

или, что то же, магнитным диполем с моментом  $m = -l/i\omega$ , помещенным в точку  $A$ :

$$j_{e2} = 0, \quad j_{m2} = l \delta(r - r_A), \quad E_2 = E_{om}, \quad H_2 = H_{om}.$$

Мы получаем тогда из (17.2)

$$H_l(r_A) = - \int_V \{E_{om}(r, r_A, l) j_e(r) - H_{om}(r, r_A, l) j_m(r)\} d^3r. \quad (17.5)$$

Линейные относительно  $j_e$  и  $j_m$  выражения (17.4) и (17.5) позволяют получать теперь любые моменты напряженностей флуктуационного поля, если нам известны моменты того же порядка для токов  $j_e$  и  $j_m$ . В частности, мы заведомо можем вычислить вторые моменты компонент флуктуационного поля, поскольку вторые моменты сторонних токов дает нам (16.4).

Найдем, например, среднее значение произведения  $E_{l_1}(r_1) H_{l_2}^*(r_2)$ . Полагая в (17.4)  $l = l_1$ ,  $r_A = r_1$  и записывая скалярные произведения векторов в компонентах ( $ab = a_\alpha b_\alpha$ , где по дважды вхо-

дящим индексам производится суммирование от 1 до 3), имеем

$$E_{i_1}(\mathbf{r}_1) = \int \{E_{0e\alpha}(\mathbf{r}, \mathbf{r}_1, l_1) j_{e\alpha}(\mathbf{r}) - H_{0e\alpha}(\mathbf{r}, \mathbf{r}_1, l_1) j_{m\alpha}(\mathbf{r})\} d^3r. \quad (17.6)$$

Полагая в (17.5)  $l = l_2$ ,  $\mathbf{r}_A = \mathbf{r}_2$ , получаем

$$H_{i_2}(\mathbf{r}_2) = - \int \{E_{0m\beta}(\mathbf{r}', \mathbf{r}_2, l_2) j_{e\beta}(\mathbf{r}') - H_{0m\beta}(\mathbf{r}', \mathbf{r}_2, l_2) j_{m\beta}(\mathbf{r}')\} d^3r'. \quad (17.7)$$

Для краткости мы далее опустим аргументы  $\mathbf{r}_1, l_1$  в  $E_{0e}$ ,  $H_{0e}$  и аргументы  $\mathbf{r}_2, l_2$  в  $E_{0m}$ ,  $H_{0m}$ . Умножив (17.6) на выражение, комплексно сопряженное (17.7), и взяв среднее по равновесному ансамблю случайных источников  $j_e$  и  $j_m$ , находим

$$\begin{aligned} \langle E_{i_1}(\mathbf{r}_1) H_{i_2}(\mathbf{r}_2) \rangle &= - \int \int \{E_{0e\alpha}(\mathbf{r}) E_{0m\beta}^*(\mathbf{r}') \langle j_{e\alpha}(\mathbf{r}) j_{e\beta}^*(\mathbf{r}') \rangle + \\ &+ H_{0e\alpha}(\mathbf{r}) H_{0m\beta}^*(\mathbf{r}') \langle j_{m\alpha}(\mathbf{r}) j_{m\beta}^*(\mathbf{r}') \rangle - \\ &- E_{0e\alpha}(\mathbf{r}) H_{0m\beta}^*(\mathbf{r}') \langle j_{e\alpha}(\mathbf{r}) j_{m\beta}^*(\mathbf{r}') \rangle - \\ &- H_{0e\alpha}(\mathbf{r}) E_{0m\beta}^*(\mathbf{r}') \langle j_{m\alpha}(\mathbf{r}) j_{e\beta}^*(\mathbf{r}') \rangle\} d^3r d^3r'. \end{aligned}$$

Согласно (16.4) последние два члена интегранда равны нулю в силу некоррелированности электрических и магнитных сторонних токов; корреляционные же функции токов в первых двух членах содержат  $\delta(\mathbf{r} - \mathbf{r}')$ , что позволяет сразу выполнить интегрирование по  $\mathbf{r}'$ . В результате

$$\langle E_{i_1}(\mathbf{r}_1) H_{i_2}(\mathbf{r}_2) \rangle = \frac{i\omega\Theta(\omega, T)}{8\pi^2} \int \{E_{0e\alpha} E_{0m\beta}^* (\epsilon_{\alpha\beta} - \epsilon_{\beta\alpha}^*) + H_{0e\alpha} H_{0m\beta}^* (\mu_{\alpha\beta} - \mu_{\beta\alpha}^*)\} d^3r. \quad (17.8)$$

Но величина

$$q_{em}(\mathbf{r}) = - \frac{i\omega}{16\pi} \{E_{0e\alpha} E_{0m\beta}^* (\epsilon_{\alpha\beta} - \epsilon_{\beta\alpha}^*) + H_{0e\alpha} H_{0m\beta}^* (\mu_{\alpha\beta} - \mu_{\beta\alpha}^*)\} \quad (17.9)$$

представляет собой не что иное, как спектральную объемную плотность смешанных тепловых потерь (электрических + магнитных) в точке  $\mathbf{r}$ <sup>1)</sup>. Интеграл от  $q_{em}(\mathbf{r})$  по  $\mathbf{r}$  — это спектральная

<sup>1)</sup> Поясним понятие «смешанных» потерь на примере изотропной среды, обладающей только электрическими потерями, когда  $\epsilon_{\alpha\beta} = \left(\epsilon' + i \frac{4\pi\sigma}{\omega}\right) \delta_{\alpha\beta}$  и  $\mu_{\alpha\beta} = \mu' \delta_{\alpha\beta}$ , где  $\epsilon'$  и  $\mu'$  — вещественные проницаемости,  $\sigma$  — электрическая проводимость. Формула (17.9) принимает для такой среды вид

$$q_{em}(\mathbf{r}) = 1/2 \sigma E_{0e} E_{0e}^*, \quad q_{em} = (q_{em})^*.$$

См. продолжение примечания на следующей странице.)

плотность потерь во всем объеме среды (фактически во всех областях, где есть диссипация энергии, т. е.  $\epsilon_{\alpha\beta} - \epsilon_{\beta\alpha}^* \neq 0$  и (или)  $\mu_{\alpha\beta} - \mu_{\beta\alpha}^* \neq 0$ ):

$$Q_{em^*}(r_1, l_1; r_2, l_2) = \int q_{em^*}(r) d^3r. \quad (17.10)$$

Конечно,  $q_{em^*}$  и  $Q_{em^*}$  зависят от координат и направлений точечных источников вспомогательных дифракционных полей, в данном случае — электрического диполя с моментом  $p = -l_1/i\omega$  в точке  $r_1$  и магнитного диполя с моментом  $m = -l_2/i\omega$  в точке  $r_2$ . Из (17.8) — (17.10) следует, что

$$\langle E_{l_1}(r_1) H_{l_2}^*(r_2) \rangle = -\frac{2}{\pi} \Theta(\omega, T) Q_{em^*}(r_1, l_1; r_2, l_2). \quad (17.11)$$

Аналогичный расчет при помощи формул (17.4), (17.5) и (16.4) функций корреляции компонент  $E$  или компонент  $H$  приводит к следующим результатам:

$$\langle E_{l_1}(r_1) E_{l_2}^*(r_2) \rangle = \frac{2}{\pi} \Theta(\omega, T) Q_{ee^*}(r_1, l_1; r_2, l_2), \quad (17.12)$$

$$\langle H_{l_1}(r_1) H_{l_2}^*(r_2) \rangle = \frac{2}{\pi} \Theta(\omega, T) Q_{mm^*}(r_1, l_1; r_2, l_2). \quad (17.13)$$

В формулу (17.12) входят смешанные потери дифракционных полей от двух электрических диполей ( $p_1 = -l_1/i\omega$  в точке  $r_1$  и  $p_2 = -l_2/i\omega$  в  $r_2$ ), а в (17.13) — от двух магнитных диполей ( $m_1 = -l_1/i\omega$  в  $r_1$  и  $m_2 = -l_2/i\omega$  в  $r_2$ ). Не смешанные, а собственные потери дифракционного поля одного источника определяют *среднеквадратичное значение какой-либо одной* компоненты флуктуационного поля в *одной и той же* точке. Например, полагая в (17.12)  $r_1 = r_2 = r$  и  $l_1 = l_2 = l$ , получаем

$$\langle |E_l(r)|^2 \rangle = \frac{2}{\pi} \Theta(\omega, T) Q_{ee^*}(r, l; r, l), \quad (17.14)$$

Мы назвали эти потери «смешанными», так как они обусловлены полями от двух разных источников. Посмотрим на том же примере, как такие потери возникают.

Как известно, мгновенная объемная плотность джоулева тепла есть  $q(l, r) = \sigma(r) E^2(l, r)$ . Для гармонического поля  $E(l, r) = 1/2 [E(\omega, r) e^{-i\omega t} + E^*(\omega, r) e^{i\omega t}]$  в среднем по периоду  $T = 2\pi/\omega$  имеем

$$q(\omega, r) = 1/4 \sigma(r) |E(\omega, r)|^2.$$

Если поле  $E$  представляет собой суперпозицию двух полей ( $E = E_1 + E_2$ ), то, очевидно,

$$q(\omega, r) = q_{11^*} + q_{12^*} + q_{1^*2} + q_{22^*},$$

где  $q_{11^*}(\omega, r) = 1/4 \sigma |E_1|^2$ ,  $q_{22^*}(\omega, r) = 1/4 \sigma |E_2|^2$  — потери каждого из полей  $E_1$  и  $E_2$  в отдельности, а  $q_{12^*}(\omega, r) = 1/4 \sigma E_1 E_2^*$ ,  $q_{1^*2}(\omega, r) = 1/4 \sigma E_1^* E_2$  — смешанные потери полей  $E_1$  и  $E_2$ .

т. е. сюда входят собственные потери дифракционного поля, создаваемого электрическим диполем с моментом  $p = -1/i\omega$ , находящимся в точке  $g$ .

Формулы (17.11)—(17.13) можно назвать *кирхгофовой формой* ФДТ, потому что они представляют собой *прямое обобщение закона Кирхгофа* классической теории теплового излучения. Этот закон связывает, как известно, интенсивность теплового излучения тела в каком-либо направлении с поглощением в этом теле при падении на него плоской волны с обратным направлением распространения (можно сказать, что это волна от бесконечно удаленного точечного источника). Обобщение касается сразу трех сторон дела.

Во-первых, мы можем теперь находить средние значения произведений *любых* компонент  $E$  и  $H$ , а не только тех, которые определяют плотность энергии и ее поток (поток вектора Пойнтинга)—две величины, которыми только и интересуется классическая теория излучения.

Во-вторых, мы можем вычислять не только средние произведения компонент, взятых в *одной и той же* точке ( $r_1 = r_2$ ), что требуется для расчета плотности энергии и ее потока, но и пространственные *функции корреляции* флуктуационного поля ( $r_1 \neq r_2$ ).

И в-третьих, что наиболее существенно, в формулах (17.11)—(17.13) нет никаких ограничений для соотношения между длиной волны  $\lambda$  и характерными масштабами задачи  $l$  (размерами тел, радиусами кривизны их поверхностей, расстояниями от тела до точки наблюдения и т. п.). Иначе говоря, в отличие от классической теории теплового излучения, связанной условиями применимости геометрической оптики, мы можем вычислять теперь вторые моменты флуктуационного поля—как волновой его части (с учетом всех дифракционных явлений), так и неволновой (квазистационарной)—при любом соотношении  $\lambda$  и  $l$ .

Необходимо остановиться на преимуществах, которые дает обобщенный закон Кирхгофа и в отношении вычислительной стороны дела. Конечно, для нахождения вспомогательных дифракционных полей (функций Грина) по-прежнему необходимо решать обычными методами соответствующие краевые электродинамические задачи. Однако эти задачи проще тех, о которых говорилось в конце предыдущего параграфа. Для формул (17.11)—(17.13) надо находить решения однородных уравнений Максвелла, обладающие дипольными особенностями в заданных точках  $g_1$  и  $g_2$ , а не решения неоднородных уравнений с заданным, но произвольным распределением сторонних токов  $j_e$  и  $j_m$ . Более того, во многих случаях можно использовать уже известные функции Грина, т. е. готовые решения задач с точечными источниками (например, классической задачи А. Зоммерфельда о поле

диполя, расположенного над плоской границей поглощающей среды).

Наконец, далеко не всегда необходимо вычислить *напряженности* дифракционных полей. Ведь в формулы (17.11)—(17.13) входят, в конечном счете, не эти напряженности, а *тепловые потери* дифракционных полей. Во многих практически интересных случаях эти потери можно с достаточной точностью получить *приближенными* методами. Сюда относится, например, случай хорошо проводящих тел (сильный скин-эффект), случай тел, больших по сравнению с  $\lambda$  или, наоборот, малых по сравнению с  $\lambda$  (хотя бы по некоторым своим размерам, как это имеет место для тонких проводов). Далее мы приведем несколько примеров применения обобщенного закона Кирхгофа и заодно проиллюстрируем предельный переход к результатам классической теории теплового излучения. В заключение же этого параграфа сделаем еще два общих замечания.

Первое касается распространения обобщенного закона Кирхгофа (17.11)—(17.13) на случай *неравномерно нагретых* тел. Если градиенты температуры достаточно малы, так что  $\Theta = \Theta(r)$  — настолько плавная функция, что роль неравновесных процессов еще пренебрежимо мала, то, очевидно, учет неравномерного нагрева сведется к тому, что  $\Theta(r)$  надо оставлять под интегралом по объему. Вместо произведения  $\Theta Q_0$  в формулах (17.11)—(17.13) надо при таких квазиоднородных условиях писать  $\int \Theta dQ_0$ , где  $dQ_0 = q_0 d^3r$  — тепловые потери дифракционного поля в элементе объема  $d^3r$  рассматриваемого тела.

Второе замечание касается *нулевых колебаний*. Во все формулы корреляционной теории равновесных и квазиравновесных флуктуационных полей входит множителем средняя энергия осциллятора  $\Theta(\omega, T)$ , содержащая слагаемое  $\hbar\omega/2$  — энергию так называемых *нулевых колебаний* (см. (15.10)). Слагаемое в (15.10), зависящее от температуры  $T$  и обращающееся в нуль при  $T = 0$ , соответствует так называемому *черному излучению*. Только эта часть обычно и рассматривается, когда речь идет об излучении, т. е. о *потоке* энергии, поскольку лишь она является в таких случаях наблюдаемой величиной. Между тем в формулах типа (17.11) содержится поток энергии и нулевых колебаний. Когда и почему его не следует учитывать?

Дело в том, что при выводе формул типа (17.11)—(17.13) сделано неявное допущение, что рассматриваемое тело является *единственным* источником флуктуационного поля. В действительности же нулевые колебания существуют и в отсутствие данного тела, так как они создаются *всеми* телами без исключения, в том числе и абсолютно холодными ( $T = 0$ ). Можно сказать, что по отношению к нулевым колебаниям всегда имеет место равно-



весное состояние, т. е. нулевые колебания — это всегда *стоячие волны* и, соответственно, любой поток энергии этих колебаний всюду гасится встречным потоком той же интенсивности<sup>1)</sup>. Поэтому в любых формулах, относящихся к *потоку* энергии (но не к ее *плотности*!), следует удерживать лишь ту часть  $\Theta(\omega, T)$ , которая относится к черному излучению, вычитая поток энергии нулевых колебаний, который всегда компенсирован при любом окружении данного тела.

### § 18. Примеры применений обобщенного закона Кирхгофа

Рассмотрим в качестве первого примера случай полупространства  $z < 0$ , заполненного поглощающей однородной и изотропной средой с проницаемостями  $\epsilon_1, \mu_1$ , над которым ( $z > 0$ ) среда тоже однородна и изотропна, но прозрачна (проницаемости  $\epsilon$  и  $\mu$  вещественны) (рис. 13). Как мы знаем, для нахождения средних моментов флуктуационного поля в какой-либо точке прозрачной среды надо знать дифракционные поля  $E_{oe}, H_{oe}$  и  $E_{om}, H_{om}$  элементарных электрического и магнитного диполей с моментами  $p = -l_1/i\omega$  и  $m = -l_2/i\omega$ , помещенных в эту точку ( $l_1$  и  $l_2$  — единичные векторы). Решение этой классической задачи можно найти в любом учебнике по распространению радиоволн (подробное изложение см., например, в монографиях [9, 10]). Потери дифракционного поля  $E_0 = E_{oe} + E_{om}, H_0 = H_{oe} + H_{om}$  в поглощающем полупространстве  $z < 0$  проще всего вычислить как поток энергии через границу  $z = 0$  в это полупространство:

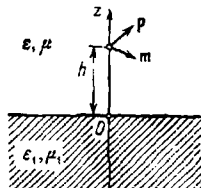


Рис. 13.

$$Q_0 = -\frac{c}{16\pi} \int_0^\infty r dr \int_0^{2\pi} d\varphi \{ [E_0 H_0]_x + [E_0 H_0]_z \}_{z=0}. \quad (18.1)$$

На плоскости  $z = 0$  мы ввели в (18.1) полярные координаты  $r$  и  $\varphi$ , причем диполи находятся на высоте  $h$  над началом отсчета  $r = 0$ . Ясно, что  $Q_0$  будет зависеть только от  $h$  и, конечно, от ориентаций диполей  $l_1$  и  $l_2$ .

Разные выражения для  $Q_0$ , отвечающие различным ориентациям  $l_1$  и  $l_2$ , определяют по формулам (17.11) — (17.14) средние значения произведений соответствующих компонент флуктуационного поля. Если, скажем, направить  $l_1$  по оси  $x$  и положить

<sup>1)</sup> Разумеется, гасится в среднем, т. е. гасятся средние значения встречающихся потоков. Если интересоваться флуктуациями потока энергии, то в них вносят вклад и нулевые колебания.

$m=0$ , то мы получим  $\langle |E_x|^2 \rangle$ ; если направить  $I_1$  по оси  $x$ , а  $I_2$  по оси  $y$ , то соответствующее  $Q_0$  даст  $\langle E_x H_y^* \rangle$  и т. д. Отличными от нуля оказываются только следующие моменты<sup>1)</sup>:

$$\begin{aligned} \langle |E_x|^2 \rangle &= \langle |E_y|^2 \rangle, \quad \langle |H_x|^2 \rangle = \langle |H_y|^2 \rangle, \\ \langle |E_z|^2 \rangle &= \langle |H_z|^2 \rangle, \quad \langle E_x H_y^* \rangle = -\langle E_y H_x^* \rangle, \end{aligned} \quad (18.2)$$

зависящие уже только от высоты  $h$  точки наблюдения над границей раздела.

Зная моменты (18.2), нетрудно составить выражения для спектральных плотностей (по  $\omega > 0$ ) электрической и магнитной энергий флуктуационного поля в прозрачной среде (с  $\epsilon$  и  $\mu$ , не зависящими от  $\omega$ , т. е. в отсутствие дисперсии) и для  $z$ -компоненты вектора Пойнтинга:

$$\begin{aligned} u_{ew}(h) &= \frac{\epsilon}{4\pi} \langle |E|^2 \rangle, \quad u_{mw}(h) = \frac{\mu}{4\pi} \langle |H|^2 \rangle, \\ \mathcal{S}_{\omega z}(h) &= \frac{c}{4\pi} \langle E_x H_y^* - E_y H_x^* + E_x^* H_y - E_y^* H_x \rangle. \end{aligned} \quad (18.3)$$

Как и моменты (18.2), величины (18.3) выражаются однократными интегралами от 0 до  $\infty$  по некоторой переменной  $p$ , причем расстояние  $h$  точки наблюдения от границы  $z=0$  входит в подынтегральные выражения только через экспоненту  $\exp[-(q+q^*)h]$ , где  $q = \sqrt{p^2 - k^2}$ , а  $k$  — волновое число в прозрачной среде ( $k = k_0 \sqrt{\epsilon \mu}$ ,  $k_0 = \omega/c$ ). Тем самым, интегралы по  $p$ , выражающие величины (18.3), распадаются на два существенно различных слагаемых. 1) Интегралы по интервалу  $p$  от 0 до  $k$ . Здесь  $q$  чисто мнимое,  $q+q^*=0$ , и эта доля в величинах (18.3) не зависит от  $h$ . Это волновая часть флуктуационного поля. 2) Интегралы по интервалу  $p$  от  $k$  до  $\infty$ , где  $q$  вещественно и положительно; соответствующие вклады в  $u_{ew}(h)$  и  $u_{mw}(h)$  убывают с удалением от границы  $z=0$ , а в потоке энергии  $\mathcal{S}_{\omega z}$  эта часть вообще отсутствует, так как в подынтегральное выражение для  $\mathcal{S}_{\omega z}$  входит еще и множитель  $q - q^*$ , который при вещественном  $q$  обращается в нуль. Это квазистационарное флуктуационное поле, не дающее вклада в поток энергии.

Для волнового поля, если в качестве переменной интегрирования ввести вместо  $p$  угол  $\theta$  между волновым вектором  $\mathbf{k}$  и плоскостью  $z=0$ , так что  $p = k \sin \theta$ , величины (18.3) приводятся

<sup>1)</sup> Подробно все вычисления приведены в [6], § 6.

к виду

$$u_{\omega\omega}^n = u_{m\omega}^n = \frac{u_{\omega}^n}{2} = \frac{u_{0\omega} n^2}{4} \int_0^{\pi/2} [1 - \mathcal{R}(\theta)] \sin \theta d\theta, \quad (18.4)$$

$$\mathcal{S}_{\omega z}^n = 2\pi \mathcal{I}_{0\omega} n^2 \int_0^{\pi/2} [1 - \mathcal{R}(\theta)] \cos \theta \sin \theta d\theta. \quad (18.5)$$

Здесь  $n = \sqrt{\varepsilon\mu}$  — показатель преломления прозрачной среды,  $u_{0\omega}$  и  $\mathcal{I}_{0\omega} = \frac{c}{4\pi} u_{0\omega}$  — соответственно плотность энергии и интенсивность равновесного излучения в *вакууме*, а  $\mathcal{R}(\theta)$  — полусумма френелевских коэффициентов отражения (по энергии) при угле падения плоской волны  $\theta$  и при двух ее поляризациях — с электрическим вектором, параллельным плоскости падения и перпендикулярным к ней:

$$\mathcal{R}(\theta) = 1/2 [\mathcal{R}_{\parallel}(\theta) + \mathcal{R}_{\perp}(\theta)], \quad (18.6)$$

$$\mathcal{R}_{\parallel}(\theta) = \left| \frac{\frac{\varepsilon_1}{\varepsilon} \cos \theta - \sqrt{N^2 - \sin^2 \theta}}{\frac{\varepsilon_1}{\varepsilon} \cos \theta + \sqrt{N^2 - \sin^2 \theta}} \right|^2,$$

$$\mathcal{R}_{\perp}(\theta) = \left| \frac{\frac{\mu_1}{\mu} \cos \theta - \sqrt{N^2 - \sin^2 \theta}}{\frac{\mu_1}{\mu} \cos \theta + \sqrt{N^2 - \sin^2 \theta}} \right|^2$$

( $N = \sqrt{\varepsilon_1 \mu_1 / \varepsilon \mu}$  — относительный показатель преломления). В (18.6) вошла полусумма  $\mathcal{R}_{\parallel}$  и  $\mathcal{R}_{\perp}$ , так как в рассматриваемой задаче обе поляризации совершенно равноправны.

Сопоставление (18.5) с общим выражением нормальной к поверхности компоненты вектора Пойнтинга  $\mathcal{S}_{\omega z}$  через интенсивность  $\mathcal{I}_{\omega}(\theta, \varphi)$ :

$$\mathcal{S}_{\omega z} = \int_{\theta < \pi/2} \mathcal{I}_{\omega}(\theta, \varphi) \cos \theta d\varphi = \int_0^{2\pi} d\varphi \int_0^{\pi/2} \mathcal{I}_{\omega}(\theta, \varphi) \cos \theta \sin \theta d\theta,$$

показывает, что в нашем случае, т. е. в случае излучения полупространства, заполненного однородной и изотропной поглощающей средой, интенсивность  $\mathcal{I}_{\omega}$  не зависит от угла  $\varphi$  и равна (закон Кирхгофа)

$$\mathcal{I}_{\omega}(\theta) = \mathcal{I}_{0\omega} n^2 [1 - \mathcal{R}(\theta)]. \quad (18.7)$$

Отсюда можно перейти к случаю бесконечного пространства, заполненного прозрачной средой с показателем преломления  $n$ , просто положив  $\mathcal{R} = 0$ . Тогда получается закон Клаузюса,

связывающий равновесные интенсивности в прозрачной среде и в вакууме:

$$\mathcal{J}_\omega = \mathcal{J}_{\omega 0} n^2.$$

Тот же переход к  $\mathcal{R} = 0$  дает, согласно (18.4),

$$u_\omega^s = \frac{1}{2} u_{\omega 0} n^2,$$

т. е. *половину* плотности энергии  $u_{\omega 0} n^2$  равновесного излучения в прозрачной среде, как и должно быть, поскольку мы выделили из этого излучения *односторонний* поток энергии (в сторону  $z > 0$ ).

Для *квазистационарного* поля квадратуры, выражающие плотности энергии  $u_{\omega 0}^{KB}$  и  $u_{m\omega}^{KB}$ , сложнее, и можно дать лишь приближенные их оценки.

Если  $kh$  гораздо больше величин  $\left| \frac{\varepsilon}{\varepsilon_1} \sqrt{N^2 - 1} \right|$  и  $\left| \frac{\mu}{\mu_1} \sqrt{N^2 - 1} \right|$ , то

$$u_{\omega 0}^{KB} \approx u_{m\omega}^{KB} \approx \frac{u_\omega^{KB}}{2} \approx \frac{u_{\omega 0} n}{8(kh)^2} \operatorname{Re} \left( \frac{\varepsilon_1 \mu + \mu_1 \varepsilon}{\sqrt{N^2 - 1}} \right),$$

т. е. убывание плотностей энергии с удалением от границы (с ростом  $h$ ) происходит по закону  $1/h^2$ . На расстояниях  $h$ , при которых

$$(kh)^2 \geq \frac{1}{4n^2} \operatorname{Re} \left( \frac{\varepsilon_1 \mu + \mu_1 \varepsilon}{\sqrt{N^2 - 1}} \right), \quad (18.8)$$

плотность энергии квазистационарной части теплового поля уже пренебрежимо мала по сравнению с постоянной в полупространстве  $z > 0$  плотностью энергии волновой части поля (18.4). Напротив, на малых расстояниях  $h$  преобладает квазистационарное поле, так что малые зазоры и полости, для которых условие (18.8) не выполнено, практически заполнены именно квазистационарным флуктуационным полем.

При малых  $kh$  с точностью до членов порядка  $1/(kh)^2$  получаются оценки

$$\begin{aligned} u_{\omega 0}^{KB} &\approx \frac{u_{\omega 0} n^2 \varepsilon}{4(kh)^2} \left\{ \frac{\varepsilon_1}{|\varepsilon + \varepsilon_1|^2 kh} + \operatorname{Re} \left( \frac{\varepsilon_1 \sqrt{N^2 - 1}}{(\varepsilon + \varepsilon_1)^2} \right) \right\}, \\ u_{m\omega}^{KB} &\approx \frac{u_{\omega 0} n^2 \mu}{4(kh)^2} \left\{ \frac{\mu_1^2}{|\mu + \mu_1|^2 kh} + \operatorname{Re} \left( \frac{\mu_1 \sqrt{N^2 - 1}}{(\mu + \mu_1)^2} \right) \right\}. \end{aligned} \quad (18.9)$$

Отсюда видно, что при наличии только электрических потерь ( $\mu_1^2 = \operatorname{Im} \mu_1 = 0$ ) плотность электрической энергии  $u_{\omega 0}^{KB}$  растет с уменьшением  $h$ , как  $1/h^2$ , а магнитной  $u_{m\omega}^{KB}$  — лишь как  $1/h^2$  (а в случае, когда  $\varepsilon_1^2 = \operatorname{Im} \varepsilon_1 = 0$ , — наоборот). Интеграл от  $u_\omega^{KB} = u_{\omega 0}^{KB} + u_{m\omega}^{KB}$  по любому конечному объему, прилегающему к границе раздела, тоже расходится. Как уже было сказано в § 16,

это результат принятой нами дельта-корреляции сторонних источников в поглощающей среде. Если бы мы взяли нелокальную среду, в которой, тем самым, сторонние флуктуационные токи обладали бы некоторым отличным от нуля радиусом корреляции  $a$ , то плотность  $u_{\text{св}}^{\text{ка}}$  оставалась бы при  $h \rightarrow 0$  конечной, достигая значения порядка  $u_{\text{св}}/(ka)^3$ .

Современная усилительная техника в диапазоне СВЧ вполне позволяет поставить прямой опыт, выявляющий сильное нарастание квазистационарного флуктуационного поля вблизи поверхности нагретого проводящего тела, но, к сожалению, такой опыт пока не осуществлен.

В рассмотренном примере мы получили для волновой части поля классический закон Кирхгофа (18.7), так как в случае бесконечного полупространства этот закон справедлив для любой длины волны. Правда, для достаточно малых расстояний  $h$  от поверхности среды были прослежены эффекты, выходящие за рамки классической теории теплового излучения (наличие квазистационарного флуктуационного поля). Но особый интерес представляют приложения общей теории к таким задачам, где классическая теория вообще неприменима, так как размеры тел не слишком велики по сравнению с длиной волны. Тогда и в волновой части теплового поля должна проявиться новая закономерность — зависимость характеристик волнового поля от отношения  $l/\lambda$ , где  $l$  — характерный размер (выше это было расстояние  $h$  от границы). Приведем пример, иллюстрирующий сказанное, — тепловое излучение равномерно нагретого шара.

Общий метод остается прежним: надо найти вспомогательное дифракционное поле диполей с моментами  $\mathbf{p} = -\mathbf{I}_1/i\omega$  и  $\mathbf{m} = -\mathbf{I}_2/i\omega$ , находящихся в какой-либо точке прозрачной среды вещественные проницаемости  $\epsilon$  и  $\mu$ ), окружающей шар, который заполнен средой с комплексными проницаемостями  $\epsilon_1$  и  $\mu_1$ . Решение этой краевой задачи (с обычными условиями непрерывности тангенциальных компонент  $\mathbf{E}_0$  и  $\mathbf{H}_0$  на поверхности шара и излучения на бесконечности) нетрудно получить в виде рядов по фундаментальным векторным функциям шара. Потери  $Q_0$  дифракционного поля в шаре можно вычислить затем как поток энергии внутрь шара через его поверхность<sup>1)</sup>. Напомним, что во всех таких, по сути дела, неравновесных задачах («высвечивание» нагретого тела) предполагается, что внутри тела каким-то образом поддерживается квазиравновесное состояние и, в частности, постоянная температура.

Мы приведем только один результат — зависимость среднего вектора Пойнтинга флуктуационного поля нагретого шара от

<sup>1)</sup> Подробное изложение расчета см. в [6], § 7, где таким же путем решена задача и для бесконечного круглого цилиндра.

радиуса шара  $a$ . На рис. 14 для случая шара, обладающего хорошей электрической проводимостью, но лишённого магнитных потерь, показана зависимость удельной мощности  $p_\omega$  (т. е. мощности, излучаемой с единицы поверхности шара) от параметра  $\alpha = k_0 a = \omega a/c = 2\pi a/\lambda$ . При  $\lambda/a \rightarrow 0$  ( $\alpha \rightarrow \infty$ ) мощность стремится к постоянному значению  $p_\infty$ , которое отвечает классическому закону Кирхгофа. По оси ординат на рис. 14 отложено отношение  $p_\omega/p_\infty$ .

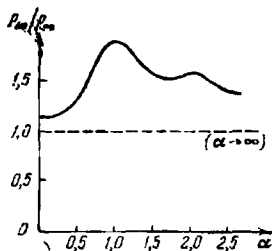


Рис. 14.

Мы видим, что с ростом  $\alpha$  удельная мощность излучения сначала растет, а затем спадает со слабо выраженными максимумами, которые несколько сдвинуты вправо от значений  $\alpha$ , соответствующих собственным частотам  $\omega$  электрических колебаний шара<sup>1)</sup>. У малого ( $\alpha \sim 1$ ) хорошо проводящего шара излучение с единицы поверхности оказывается в полтора—два раза интенсивнее, чем у большого ( $\alpha \gg 1$ ). Зависимость  $p_\omega$  от от-

ношения  $\lambda/a$ —это и есть то новое, что дает флуктуационная электродинамика.

Полный поток энергии, излучаемой шаром на частоте  $\omega$ , равен  $P_\omega = 4\pi a^2 p_\omega$  и, конечно, одинаков через любую окружающую шар замкнутую поверхность. Но средняя плотность энергии  $u_\omega$  флуктуационного поля ведет себя иначе. Вдали от шара  $u_\omega$  убывает, как  $1/R^2$ . Это плотность энергии излучения (волнового поля). При приближении же к поверхности шара  $u_\omega$  быстро нарастает и при  $R = a$  обращается в бесконечность. Разумеется, и здесь это нарастание  $u_\omega$  обусловлено квазистационарной частью флуктуационного поля.

Заметим, что решение вспомогательной дифракционной задачи о полях диполей  $p$  и  $m$  существенно облегчается в случае хорошо проводящих тел (сильно выраженный скин-эффект). Для таких тел можно воспользоваться решением дифракционной задачи с так называемыми *импедансными* граничными условиями, т. е. вообще не рассматривать поле внутри тела. В первом приближении дифракционное поле рассчитывается для идеально проводящих тел той же формы и расположения, что и рассматриваемые (кстати сказать, случай идеально проводящих тел пользуется из-за простоты граничных условий наибольшим вниманием специалистов по теории дифракции). Тепловые потери вычисляются

<sup>1)</sup> Если бы у шара были только магнитные потери, то максимумы приходились бы на собственные частоты колебаний магнитного типа. При наличии обоих видов потерь картина была бы более сложной.

затем интегрированием по поверхности тела *поверхностной плотности* джоулева тепла. На элементе  $d\Sigma$  поверхности тела поглощается тепло

$$dQ_0 = \frac{c_0^2}{8\pi} |H_{0t}|^2 d\Sigma, \quad \zeta' = \sqrt{\frac{\mu_1 \omega}{8\pi \sigma_1}}, \quad (18.10)$$

где  $\zeta'$  — вещественная часть импеданса, а  $H_{0t}$  (тангенциальная компонента  $H_0$ ) берется из решения дифракционной задачи для идеально проводящего тела.

Другой приближенный способ вычисления потерь  $Q_0$  дифракционного поля применим в тех случаях, когда, во-первых, все характерные размеры тела (включая радиусы кривизны его поверхности) велики по сравнению с длиной волны  $\lambda$  и с глубиной проникновения поля внутрь тела и, во-вторых, моменты флуктуационного поля ищутся лишь вдали от тела — на расстояниях, значительно превышающих  $\lambda$ . Соответственно точечные источники дифракционного поля тоже помещаются на таких расстояниях. Кроме того, предполагается, что форма тела такова, что в приближении геометрической оптики лучи, выходящие из точечного источника, испытывают на границе тела лишь однократные отражения.

Перечисленные условия позволяют считать, что в ближайшей окрестности каждого элемента поверхности тела поле удаленного точечного источника имеет структуру плоской волны, а значит, полное дифракционное поле у поверхности можно найти, пользуясь *френелевскими коэффициентами отражения* (для каждой из поляризаций и с локальным значением угла падения первичной волны). Так как, по предположению, многократных отражений нет, а преломленная волна поглощается, не доходя до других участков поверхности тела, потери  $Q_0$  дифракционного поля равны просто суммарной мощности преломленных в тело волн, причем интеграл берется по «освещенной» части поверхности тела.

Ряд примеров использования обоих приближенных методов расчета потерь  $Q_0$  (т. е. применение формул сильного скин-эффекта и применение френелевских отражательных формул) приведен в монографии [6].

Остановимся теперь на случае, когда излучение тела, находящегося в свободном пространстве (для простоты — в вакууме), интересует нас только в волновой зоне этого тела, т. е. на расстояниях  $R \gg l^2/\lambda$ , где  $l$  — размеры тела. Помещая диполи  $p$  и  $m$  на столь большом расстоянии, мы можем считать, что в области, занятой телом, приходящая волна — плоская. Если моменты  $p$  и  $m$  (т. е. орты  $l_1$  и  $l_2$ ) взять взаимно ортогональными и образующими с  $R$  ортогональную правовинтовую связку (рис. 15), то напряженности в падающей на тело волне будут различаться

только знаком:

$$E_{oe} = -E_{om} \equiv E_o, \quad H_{oe} = -H_{om} \equiv H_o, \quad (18.11)$$

т. е. падающая волна линейно поляризована (по  $l_1$ ), а ее амплитуда (при  $p=1/i\omega$ ) равна  $|E_o| = 2\pi/c\lambda R$ .

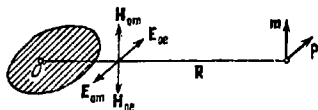


Рис. 15.

В силу (18.11) формулы (17.4) и (17.5) дают для напряженностей теплового поля в точке  $R$  выражения

$$E_1 = H_2 = \int_V (E_o j_e - H_o j_m) d^3 r, \quad (18.12)$$

где  $V$  — объем тела, а индексы 1 и 2 обозначают компоненты по ортам  $l_1$  и  $l_2$ . По формулам (17.11) — (17.13) мы получаем поэтому, что

$$\langle |E_1|^2 \rangle = \langle |H_2|^2 \rangle = \langle E_1 H_2^* \rangle = \frac{2}{\pi} \Theta(\omega, T) Q_o.$$

Соответственно поток энергии (вектор Пойтинга) излучения тела (с поляризацией  $l_1$  и в направлении  $R$ ) будет

$$\mathcal{S}_R = \frac{c}{4\pi} \langle E_1 H_2^* + E_1^* H_2 \rangle = \frac{c}{2\pi} \langle |E_1|^2 \rangle = \frac{c}{\pi^2} \Theta(\omega, T) Q_o, \quad (18.13)$$

а поток в телесный угол  $do$ , т. е. через площадку  $R^2 do$ , находящуюся на расстоянии  $R$  от тела, составит

$$dP_\omega = \mathcal{S}_R R^2 do = \frac{c}{\pi^2} \Theta(\omega, T) Q_o R^2 do. \quad (18.14)$$

Индекс  $\omega$  указывает, как обычно, что речь идет о спектре по частотам  $\omega \geq 0$ .

Выразим теперь потери  $Q_o$  дифракционного поля в теле через так называемый *эффективный поперечник поглощения*:

$$\sigma_{эфф} = \frac{Q_o}{\frac{c}{8\pi} |E_o|^2}.$$

Таким образом,  $\sigma_{эфф}$  — это та площадь фронта падающей на тело плоской волны, поток энергии через которую равен поглощаемой телом мощности  $Q_o$  (на поляризации  $l_1$ ). Подставив в  $\sigma_{эфф}$  значение  $|E_o|$ , находим

$$Q_o = \sigma_{эфф} \frac{\pi}{2c\lambda^2 R^2}.$$



Наконец, внося это выражение для  $Q_0$  в (18.14) и удвоив результат (для учета обеих независимых поляризаций), получаем

$$dP_\omega = \frac{\Theta(\omega, T)}{\pi\lambda^2} \sigma_{\text{эфф}} d\omega. \quad (18.15)$$

Таков поток энергии теплового излучения тела в дальней зоне в телесный угол  $d\omega$  в направлении  $R$  и на частоте  $\omega$ .

Если тело представляет собой перпендикулярную к  $R$  пластину, размеры которой гораздо больше  $\lambda$ , то  $\sigma_{\text{эфф}} = A\Sigma$ , где  $\Sigma$  — площадь пластины, а  $A$  — ее коэффициент поглощения. Тогда

$$dP_\omega = \frac{\Theta(\omega, T)}{\pi\lambda^2} A\Sigma d\omega = \frac{\omega^2 \Theta(\omega, T)}{4\pi^2 c^2} A\Sigma d\omega \quad (18.16)$$

в соответствии с классической теорией теплового излучения.

## § 19. Волноводная форма закона Кирхгофа

При рассмотрении передачи электромагнитных сигналов по волноводам представляет интерес спектральная интенсивность тепловых «шумов», т. е. мощность теплового излучения, переносимая по волноводу в спектральном интервале  $(\omega, \omega + d\omega)$ . Тепловое поле может создаваться стенками самого волновода, какими-либо антенными устройствами, к которым он присоединен (щель, рупор), рефлекторами, диафрагмами, аттенуаторами и т. п. Назовем для краткости любую систему таких элементов *излучателем*. Разумеется, флуктуационное *волновое* поле в волноводном режиме может быть представлено как суперпозиция *бегущих* (докристаллических) собственных волн волновода, или так называемых *собственных мод*.

На основании формулы (17.11) — (17.13) нетрудно заранее предвидеть, что мощность теплового излучения, посылаемая в волновод любым излучателем на какой-либо  $n$ -й моде и на частоте  $\omega$ , будет связана с *коэффициентом поглощения* этого излучателя, когда на него падает  $n$ -я собственная волна частоты  $\omega$ . Соответствующую формулу, названную *волноводной формой закона Кирхгофа* ([1], § 17), легко получить как при помощи теоремы взаимности, так и на основе принципа детального равновесия ([6], § 9). Мы приведем только второй способ вывода, чтобы сказать, не предполагающий заранее выполнения теоремы взаимности.

В ч. I, § 54 было показано, что произвольный излучатель, согласованный с линией (волноводом, коаксиалом) на частоте  $\omega \geq 0$ , посылает в линию в интервале частот  $(\omega, \omega + d\omega)$

мощность <sup>1)</sup>

$$P_{\omega} d\omega = \frac{\Theta(\omega, T)}{2\pi} d\omega.$$

Если согласованный излучатель возбуждает только одну ( $n$ -ю) собственную волну, то это равенство должно выполняться и для данной волны, т. е. для каждой моды и частоты, на которых излучатель согласован, имеем

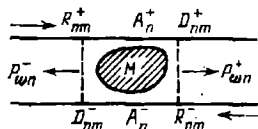


Рис. 16.

$$P_{\omega n} = \frac{\Theta}{2\pi}. \quad (19.1)$$

Рассмотрим следующую равновесную систему. Между двумя сечениями регулярного волновода, показанными пунктиром на рис. 16, расположен некоторый несогласованный излучатель

$M$  (причем рассогласование, вообще говоря, различно по обе стороны от  $M$ ).

Введем для  $M$  коэффициенты поглощения  $n$ -й волны  $A_n^{\pm}$  и коэффициенты ее трансформации ( $n \rightarrow m$ ) при отражении от  $M$  ( $R_{nm}^{\pm}$ ) и при прохождении через занятую  $M$  область ( $D_{nm}^{\pm}$ ), где индексы *плюс* и *минус* означают падение первичной  $n$ -й волны на  $M$  соответственно *слева* и *справа*. Из закона сохранения энергии имеем

$$\begin{aligned} 1 &= \sum'_m (R_{nm}^+ + D_{nm}^+) + A_n^+, \\ 1 &= \sum'_m (R_{nm}^- + D_{nm}^-) + A_n^-. \end{aligned} \quad (19.2)$$

Штрих около знака суммы должен напоминать, что суммирование распространяется только на те  $m$ -е волны, которые на частоте  $\omega$  являются бегущими (докритическими).

Пусть по обе стороны от излучателя  $M$  находятся черные излучатели, т. е. тела, которые имеют ту же температуру, что и  $M$ , и согласованы с волноводом (при данной частоте  $\omega$ ) на каждой из рассматриваемых бегущих волн. Обозначим через  $P_{\omega n}^{\pm}$  мощности, излучаемые  $M$  на  $n$ -й волне вправо и влево (рис. 16). Принцип детального равновесия требует, чтобы встречные потоки энергии на этой волне были одинаковы (как справа, так и слева

<sup>1)</sup> Разумеется, под  $\Theta(\omega, T)$  надо понимать здесь только энергию черного излучения (§ 17). Если же речь идет о некантовой области  $\hbar\omega \ll kT$ , то  $\Theta(\omega, T) = kT$  без каких-либо оговорок.

от  $M$ ):

$$\begin{aligned}\frac{\Theta}{2\pi} &= P_{\omega n}^+ + \frac{\Theta}{2\pi} \sum_m' (R_{mn}^- + D_{mn}^+), \\ \frac{\Theta}{2\pi} &= P_{\omega n}^- + \frac{\Theta}{2\pi} \sum_m' (R_{mn}^+ + D_{mn}^-).\end{aligned}\quad (19.3)$$

Умножая (19.2) на  $\Theta/2\pi$  и вычитая из соответствующих равенств (19.3), получаем

$$\begin{aligned}P_{\omega n}^+ &= \frac{\Theta}{2\pi} \left\{ A_n^- + \sum_m' (R_{nm}^- - R_{mn}^- + D_{nm}^- - D_{mn}^+) \right\}, \\ P_{\omega n}^- &= \frac{\Theta}{2\pi} \left\{ A_n^+ + \sum_m' (R_{nm}^+ - R_{mn}^+ + D_{nm}^+ - D_{mn}^-) \right\},\end{aligned}\quad (19.4)$$

что и является искомым результатом.

При выполнении принципа взаимности матрицы  $R$  и  $D$  обладают симметрией:

$$R_{nm}^\pm = R_{mn}^\pm, \quad D_{nm}^\pm = D_{mn}^\mp,$$

и мы получаем тогда из (19.4)

$$P_{\omega n}^+ = \frac{\Theta}{2\pi} A_n^-, \quad P_{\omega n}^- = \frac{\Theta}{2\pi} A_n^+, \quad (19.5)$$

т. е. мощность, излучаемая  $M$  направо (налево), определяется коэффициентом поглощения  $M$  при падении  $n$ -й волны справа (слева).

В отсутствие трансформации типов воли, когда матрицы  $R$  и  $D$  диагональны, получаем из (19.4)

$$\begin{aligned}P_{\omega n}^+ &= \frac{\Theta}{2\pi} (A_n^- + D_n^- - D_n^+), \\ P_{\omega n}^- &= \frac{\Theta}{2\pi} (A_n^+ + D_n^+ - D_n^-).\end{aligned}\quad (19.6)$$

Выполнение принципа взаимности ( $D_n^+ = D_n^-$ ) или же полная непрозрачность излучателя  $M$  ( $D_n^\pm = 0$ ) снова приводят к (19.5).

Заметим, что поскольку  $P_{\omega n}^\pm$  — неотрицательные величины, из (19.6) следует, что для любого тела или системы  $M$

$$1 - R_n^\mp = A_n^\mp + D_n^\mp \geq D_n^\pm,$$

т. е. доля энергии, поглощенная и пропущенная при облучении  $M$  в одном направлении, всегда не меньше, чем доля энергии, прошедшая через  $M$  при облучении во встречном направлении.

Если материалы уравнения для тел, составляющих излучатель  $M$ , локальны, то можно перейти к квазиравновесному случаю, считая температуру  $T$  функцией точки, и пользоваться средним значением  $\bar{\Theta}(\omega, T)$ , взвешенным по локальной величине джоулевых потерь внутри тел:

$$\bar{\Theta} = \frac{1}{Q_{0n}(\omega)} \int \Theta(\omega, T) dQ_{0n},$$

тогда (19.5) запишется в более общем виде:

$$P_{\omega n}^{\pm} = \frac{\bar{\Theta}}{2\pi} A_n^{\mp}(\omega). \quad (19.7)$$

Значения величины  $A_n^{\pm}(\omega)$  определяются, конечно, всей структурой поля, создаваемого в излучателе  $M$  падающей на него  $n$ -й волной, а не только его поглощающими элементами (см. задачу 4).

Согласно (19.5) какое-либо удовлетворяющее принципу взаимности тело излучает, находясь в волноводе, следующие *полные* мощности в интервале частот  $d\omega$ :

Согласно (19.5) какое-либо удовлетворяющее принципу взаимности тело излучает, находясь в волноводе, следующие *полные* мощности в интервале частот  $d\omega$ :

$$P_{\omega}^{\pm} d\omega = \frac{\bar{\Theta}}{2\pi} \sum_n' A_n^{\mp}(\omega) d\omega. \quad (19.8)$$

Сумма берется по всем докритическим волнам, число которых

при разных значениях  $\omega$  различно. Рис. 17 иллюстрирует формулу (19.8) на примере одностороннего излучения хорошо проводящей поперечной перегородки в прямоугольном волноводе. Мощность  $P_{\omega}$  (в некотором произвольном масштабе) показана в функции параметра  $\xi = 2a/\lambda$ , где  $a$  — меньшая сторона прямоугольного сечения. С ростом  $\xi$  (повышением  $\omega$ ) все большее число собственных  $E$ - и  $H$ -волн переходит в разряд бегущих. Из-за того, что перегородка обладает только электрическими потерями и лишена магнитных, коэффициенты поглощения  $E$ - и  $H$ -волн ведут себя при переходе через критические частоты различно: для  $H$ -волн они плавно нарастают (от нуля при критической частоте), а для  $E$ -волн они начинаются с острых пиков, расположенных вплотную к критической частоте. Эти пики и дают сильные выбросы  $P_{\omega}$ , которые в принятом масштабе далеко выходят за пределы чертежа. Пунктиром на рисунке показан тот ход  $P_{\omega}(\xi)$ , который получился бы при экстраполяции на область малых  $\xi$  классического закона Кирхгофа. Неприменимость этого асимптотического закона вполне очевидна.

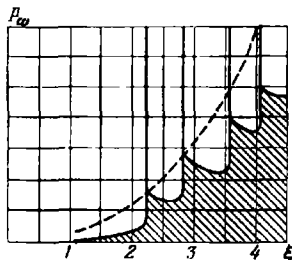


Рис. 17.

Следует отметить, что для любой совокупности волн, распространяющихся в *одном измерении* (например, по радиальному направлению в случае излучения шара), можно представить излучаемую мощность в виде (19.8), т. е. в виде суммы по взаимно ортогональным модам. Таким образом, называя формулы типа (19.8) *волноводной* формой закона Кирхгофа, мы несколько сужаем область их применимости. Однако это оправдано, так как именно волноводы дают реальную возможность выделять волны отдельных типов ( $E$  и  $H$ ) и номеров, тогда как при излучении тел в свободное пространство представляет интерес лишь вся входящая в (19.8) сумма, причем сумма бесконечная (ввиду отсутствия критических частот).

Волноводная форма закона Кирхгофа справедлива не только в бесконечном регулярном волноводе, но и в том случае, когда излучатель (система тел)  $M$  находится на конце *полубесконечного* волновода. Пусть, например, к концу волновода присоединен рупор, перед которым и (или) внутри него расположены любые тела  $M_\alpha$  ( $\alpha = 1, 2, \dots$ ) с температурами  $T_\alpha$  (рис. 18). Под излучателем  $M$  надо понимать всю систему тел  $M_\alpha$ , заключенную внутри замкнутой поверхности, охватывающей левое полупространство и достаточно удаленной от конца волновода, т. е. пересекающей его там, где уже существует одномерная совокупность взаимно ортогональных собственных волн регулярного волновода (сечение  $\Sigma$  на рис. 18). Мощность теплового излучения, посылаемого системой тел  $M$  в волновод на  $n$ -й собственной волне, запишется в том же виде (19.7):

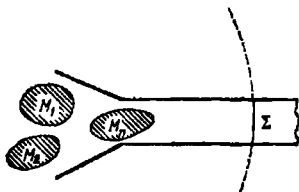


Рис. 18.

$$P_{\omega n}^+ = \frac{1}{2\pi} \sum_{\alpha} \Theta_{\alpha} A_{\alpha n}^-(\omega), \quad (19.9)$$

где суммирование распространяется на все тела  $M_\alpha$  (включая, конечно, и рупор, если в нем есть джоулевы потери).

Напомним, что  $A_{\alpha n}^-(\omega)$  — это поглощенная телом  $M_\alpha$  доля мощности  $n$ -й волны, падающей справа через сечение  $\Sigma$ , т. е. при работе всего устройства в *передающем* режиме. При этом часть мощности излучается в окружающее пространство. Простое рассуждение показывает, однако, что можно и эту часть включить в (19.9). Действительно, допустим, что среда, заполняющая пространство, охваченное замкнутой поверхностью, и окружающая тела  $M_\alpha$ , обладает *сколь угодно малым поглощением*. Если при этом вся мощность, излучаемая в передающем режиме, в конечном счете поглощается средой, то последнюю можно рассматри-

вать как *одно из тел*  $M_\alpha$ . Таким образом, один из членов суммы (19.9) будет относиться к *среде*, а роль соответствующего  $A_n^-$  будет играть так называемый *энергетический коэффициент излучения*, т. е. доля мощности падающей справа  $n$ -й волны, *излученная* рассматриваемой системой в *непоглощающую* среду.

## § 20. Тепловое излучение и антенны

Как известно, существуют разные виды антенн, например: *тонкие* (проволочные) антенны, широко применяемые в радиовещании и связи на коротких и более длинных волнах, *зеркальные* антенны, используемые в радиоастрономии и для связи на ультракоротких и еще более коротких волнах, и др. В любой антенне происходят тепловые флуктуации в материале антенны и возникает обусловленное этими флуктуациями собственное тепловое излучение антенны. Но во многих случаях представляют интерес не только эти собственные шумы и излучение, но и воздействие на антенну (а значит, и на последующие каналы) флуктуационных полей внешних нагретых тел и сред. Мы остановимся на этих *наведенных* тепловых шумах и только на *тонких* антеннах<sup>1)</sup>.

Внешнее флуктуационное поле наводит в проводах антенны токи, которые и представляют основной интерес, так как обычно именно они являются здесь непосредственно измеряемыми величинами. В принципе, если корреляционная функция падающего на антенну теплового излучения известна, моменты наведенного тока можно найти, используя формулы теории возбуждения антенн. Однако такого рода расчеты связаны с трудностями, которые обусловлены в первую очередь тем, что флуктуационное поле не обладает дельта-корреляцией вдоль проводов антенны. Эти трудности можно в значительной мере обойти, если вновь воспользоваться функцией Грина и теоремой взаимности.

В теории тонких антенн удобно и принято оперировать не со сторонними токами, а со сторонними полями и э. д. с. Целесообразно поэтому соответствующим образом видоизменить и теорему взаимности. Запишем «электрическую часть» этой теоремы (17.2):

$$\int \mathbf{E} \mathbf{j}_e d^3r = \int \mathbf{E}_0 \mathbf{j}_e d^3r, \quad (20.1)$$

и заменим сторонние токи  $\mathbf{j}_e$  и  $\mathbf{j}_{0e}$  — *источники* флуктуационного ( $\mathbf{E}$ ) и вспомогательного дифракционного ( $\mathbf{E}_0$ ) полей на — *напряжен-*

<sup>1)</sup> Теория, позволяющая рассчитывать собственное тепловое излучение антенн, развита в книге [6], в §§ 10 и 11 — для зеркальных антенн, а в § 12 — специально для тонких антенн. § 13 указанной книги посвящен наведенным тепловым шумам в антеннах и содержит более обширный материал, чем приводимый ниже в данном параграфе.

ности сторонних полей соответственно  $K$  и  $K_0$ . Так как в спектральном представлении  $j_e = -i\omega \mathfrak{D}/4\pi$ , а  $\mathfrak{D} = \epsilon K$ , где  $\epsilon = \epsilon' + i \frac{4\pi\sigma}{\omega}$  ( $\epsilon'$  — вещественная проницаемость,  $\sigma$  — проводимость), имеем

$$j_e = -\frac{i\omega\epsilon}{4\pi} K, \quad j_{0e} = -\frac{i\omega\epsilon}{4\pi} K_0, \quad (20.2)$$

где вторая формула — для дифракционного тока  $j_{0e}$ , — разумеется, аналогична первой. Если обозначить плотности токов, *наведенных* в антенне полями  $E$  и  $E_0$ , соответственно через  $j$  и  $j_0$ , то и для них будет

$$j = -\frac{i\omega\epsilon}{4\pi} E, \quad j_0 = -\frac{i\omega\epsilon}{4\pi} E_0. \quad (20.3)$$

Из (20.2) и (20.3) следует, что  $E j_{0e} = K_0 j$  и  $E_0 j_e = K j_0$ , так что теорема (20.1) принимает вид<sup>1)</sup>

$$\int K_0(r') j(r') d^3r' = \int K(r) j_0(r) d^3r. \quad (20.4)$$

Именно такая форма теоремы взаимности — с напряженностями сторонних полей и плотностями наведенных токов — чаще всего используется в случае тонких проводов (в том числе в теории проволочных антенн), а также в квазистационарной области (в том числе в теории цепей с сосредоточенными параметрами).

При применениях (20.4) надо знать корреляционную функцию компонент стороннего флуктуационного поля  $K(r)$ . Ее нетрудно получить из первого равенства (20.2) и первой формулы ФДТ (16.4), причем мы ограничимся случаем изотропного материала проводов, когда  $\epsilon_{\alpha\beta} = \epsilon \delta_{\alpha\beta}$ :

$$\begin{aligned} \langle K_\alpha(r) K_\beta^*(r') \rangle &= \frac{16\pi^2}{\omega^2 |\epsilon|^2} \langle j_{e\alpha}(r) j_{e\beta}^*(r') \rangle = \\ &= -\frac{2i}{\omega} \Theta(\omega, T) \left( \frac{1}{\epsilon^*} - \frac{1}{\epsilon} \right) \delta_{\alpha\beta} \delta(r-r'). \end{aligned} \quad (20.5)$$

Более того, если речь идет, как в нашем случае, о металлических проволоках, то обычно можно пренебречь токами смещения внутри этих проволок и считать, что  $\epsilon \approx i \frac{4\pi\sigma}{\omega}$ . Тогда

$$\langle K_\alpha(r) K_\beta^*(r') \rangle = \frac{\Theta(\omega, T)}{\pi\sigma} \delta_{\alpha\beta} \delta(r-r'). \quad (20.6)$$

Пусть нас интересует наведенный флуктуационный ток  $I$  в некотором сечении какого-либо из проводов, входящих в со-

<sup>1)</sup> Для последующего нам удобнее обозначить переменную интегрирования в левой части иначе, чем в правой; поэтому мы и ввели  $r'$ .

став антенны. Сечение  $\Sigma$  этого тонкого (квазилинейного, как говорят в электродинамике) проводника, вообще говоря, может меняться по длине провода  $s$ , но, по условию, всюду имеет линейные размеры, малые по сравнению с длиной волны  $\lambda$  в окружающей непроводящей среде. Кроме координаты  $s$ , отсчитываемой по оси провода, введем еще радиус-вектор  $\rho$ , лежащий в плоскости поперечного сечения провода. Таким образом,  $d^3r' = d^2\rho' ds'$ .

Для того чтобы выразить ток  $I(s)$  через  $K$ , допустим, что вспомогательное стороннее поле  $K_0(r')$  действует только в поперечном сечении провода с координатой  $s' = s$ , на всем сечении постоянно и направлено по оси провода (орт  $s$ ):

$$K_0(\rho', s') = \mathcal{E}_0 s \delta(s' - s). \quad (20.7)$$

Очевидно,  $\mathcal{E}_0$  — это интегральная сторонняя э. д. с., приложенная в сечении с координатой  $s$ :

$$\int K_0 s ds' = \mathcal{E}_0.$$

Коль скоро  $K_0$  имеет вид (20.7), интеграл в левой части (20.4) равен

$$\begin{aligned} \int K_0(\rho', s') j(\rho', s') d^3\rho' ds' &= \\ &= \mathcal{E}_0 \int_{\Sigma} d^2\rho' \int j_s(\rho', s') \delta(s' - s) ds' = \\ &= \mathcal{E}_0 \int_{\Sigma} d^2\rho' j_s(\rho', s) = \mathcal{E}_0 I(s) \end{aligned}$$

т. теорема взаимности (20.4) дает

$$\mathcal{E}_0 I(s) = \int K(r) j_0(r) d^3r. \quad (20.8)$$

При помощи этой формулы нетрудно получить теперь средний квадрат модуля флуктуационного тока  $I(s)$ . Умножая (20.8) на комплексно сопряженное равенство и усредняя, получаем

$$|\mathcal{E}_0|^2 \langle |I(s)|^2 \rangle = \iint \langle K_\alpha(r) K_\beta^*(r') \rangle j_{0\alpha}(r) j_{0\beta}^*(r') d^3r d^3r'.$$

Подставив сюда функцию корреляции (20.6) стороннего теплового поля  $K$ , находим

$$|\mathcal{E}_0|^2 \langle |I(s)|^2 \rangle = \frac{2}{\pi} \int \Theta(\omega, T) dQ_0, \quad (20.9)$$

где

$$dQ_0 = \frac{\sigma}{2} |E_s|^2 d^3r = \frac{|j_s|^2}{2\sigma} d^3r \quad (20.10)$$

— джоулевы потери тока, вызываемые сосредоточенной в сечении  $s$  вспомогательной э. д. с.  $\mathcal{E}_0$ . Очевидно, равенство (20.9)



представляет собой модифицированную форму обобщенного закона Кирхгофа (17.14). Оно позволяет не решать неоднородную краевую задачу о флуктуационном поле  $E$  в материале антенны, а сводит нахождение  $\langle |I(s)|^2 \rangle$  к вычислению квадратуры, если известно решение задачи о вспомогательном поле  $E_0$  дельта-запитываемой антенны, т. е. о поле, возбуждаемом э. д. с.  $\mathcal{E}_0$ , приложенной в интересующем нас сечении  $s$ .

Если  $Z_{вх}(s)$  — входной импеданс антенны по отношению к э. д. с., включенной в сечение  $s$ , то  $\mathcal{E}_0 = Z_{вх}(s) I_0(s)$ , где

$$I_0(s) = \int j_{0s}(\rho, s) d^2\rho$$

— полная сила «дифракционного» тока в этом сечении. Можно ввести аналогичным образом эквивалентную тепловую э. д. с.  $\hat{\mathcal{E}}(s)$  в сечении  $s$ , определив ее через импеданс  $Z_{вх}(s)$  и полный флуктуационный ток:

$$\hat{\mathcal{E}}(s) = Z_{вх}(s) I(s). \quad (20.11)$$

Таким образом,  $\hat{\mathcal{E}}(s) = \mathcal{E}_0 I(s) / I_0(s)$ , т. е., согласно (20.8),

$$\hat{\mathcal{E}}(s) = \frac{1}{I_0(s)} \int K(r) j_0(r) d^2r. \quad (20.12)$$

Таково выражение эквивалентной тепловой э. д. с. в сечении  $s$  в общем случае, когда обуславливающее ее стороннее поле  $K(r)$  произвольно распределено в объеме проводов антенны. В силу (20.9) средний квадрат модуля  $\hat{\mathcal{E}}(s)$  равен

$$\langle |\hat{\mathcal{E}}(s)|^2 \rangle = \frac{2}{\pi |I_0(s)|^2} \int \Theta(\omega, T) dQ_0. \quad (20.13)$$

Следует подчеркнуть, что  $\hat{\mathcal{E}}(s)$  не представляет собой флуктуационной э. д. с., распределенной вдоль провода. Формула (20.11) вводит (по сути дела, формально) для каждого сечения  $s$  свою сосредоточенную в этом сечении э. д. с.  $\hat{\mathcal{E}}(s)$ , смысл которой состоит только в том, что она обеспечивает правильное значение флуктуационного тока в том же сечении.

Однако в случае квазистационарной цепи, когда в каждой ее ветви полная сила тока  $I$  и входной импеданс цепи  $Z_{вх}$  не зависят от  $s$ , э. д. с.  $\hat{\mathcal{E}}$  тоже не зависит от  $s$ , т. е. может быть включена в любое сечение данной ветви. Если к тому же цепь имеет всюду одинаковую температуру ( $\Theta = \text{const}$ ), то  $\hat{\mathcal{E}}$  можно отождествить с локальной найквистовской э. д. с. е того двухполосника, который получается при размыкании рассматриваемой ветви цепи. При этих условиях (квазистационарности и равнове-

ности) (20.13) переходит в формулу Найквиста:

$$\langle |\tilde{e}(\omega)|^2 \rangle = \frac{2}{\pi} \Theta \frac{Q_0}{|I_0|^2} = \frac{1}{\pi} \Theta R_0, \quad (20.14)$$

где  $R_0$  — энергетическое (активное) сопротивление двухполюсника:

$$Q_0 = \frac{1}{2} R_0 |I_0|^2.$$

Вернемся к антенной формуле (20.13) и рассмотрим два иллюстрирующих ее примера.

1. Антенна в поле равновесного излучения. Пусть антенна находится в свободном пространстве, заполненном прозрачной средой со всюду одинаковой температурой ( $\Theta = \text{const}$ ).

Включенная в каком-то сечении  $s$  антенны вспомогательная э. д. с.  $\mathcal{E}_0$  создает дифракционное поле антенны, которое отнюдь не локализовано только в проводах антенны или в их ближайшей окрестности. Напротив, это поле содержит и излучаемые антенной волны, так что джоулевы потери  $Q_0$  имеют место как в материале самой антенны, так и в любых проводящих телах, оказавшихся на пути излучаемых волн. Пока мы предположим, что таких тел нет (свободное пространство) и, следовательно, мощность, отдаваемая э. д. с.  $\mathcal{E}_0$ , расходуется только на нагревание самой антенны и на излучение, причем во всякой антенне, отвечающей своему назначению, подавляющая доля приходится именно на излучение.

Предполагая, что среда обладает исчезающе малой проводимостью, и пренебрегая тепловыми потерями в самой антенне, мы можем считать, что полные потери  $Q_0$  дифракционного поля просто равны *мощности излучения*  $P_0$ . Последняя записывается обычно в виде  $P_0 = R_\Sigma(s) |I_0(s)|^2/2$ , где  $R_\Sigma$  — так называемое *сопротивление излучения* антенны, зависящее, конечно, от того, в каком сечении  $s$  включена э. д. с.  $\mathcal{E}_0$ . Таким образом,

$$Q_0 = \frac{1}{2} R_\Sigma(s) |I_0(s)|^2. \quad (20.15)$$

При  $\Theta = \text{const}$  формула (20.13) имеет вид  $\langle |\hat{\mathcal{E}}(s)|^2 \rangle = 2\Theta Q_0/\pi |I_0(s)|^2$ . Подставив сюда (20.15), получаем

$$\langle |\hat{\mathcal{E}}(s)|^2 \rangle = \frac{1}{\pi} \Theta R_\Sigma(s),$$

а спектральная плотность флуктуационной э. д. с. по положительным частотам вдвое больше:

$$\mathcal{E}_0^2(s) = 2 \langle |\hat{\mathcal{E}}(s)|^2 \rangle = \frac{2}{\pi} \Theta R_\Sigma(s). \quad (20.16)$$

Из вывода ясно, что этот результат справедлив для любой сколь угодно сложной антенны в свободном пространстве, если только отнести  $R_\Sigma$  и  $\hat{\sigma}_\omega^2$  к одному и тому же сечению антенны.

2. Тепловые шумы, наводимые удаленными телами. Если тело находится на достаточно большом расстоянии  $R$  от антенны, точнее — в ее френгоферовой зоне, и размеры тела тоже малы по сравнению с зоной Френеля, то в области пространства, занятой телом, можно считать антенное поле суперпозиций плоской волны и дифрагированного телом вторичного поля. Обозначая, как и в § 18, эффективный поперечник поглощения тела для падающей на него плоской волны соответствующей поляризации через  $\sigma_{\text{эфф}}$ , можно записать поглощаемую телом мощность в виде  $Q_0 = \sigma_{\text{эфф}} \mathcal{S}_0$ , где  $\mathcal{S}_0$  — модуль вектора Пойнтинга первичной волны. Связь  $\mathcal{S}_0$  с полной мощностью  $P_0 = R_\Sigma |I_0|^2/2$ , излучаемой антенной, дается известным соотношением

$$\mathcal{S}_0 = P_0 \frac{G}{4\pi R^2},$$

где  $G$  — функция направления, называемая *приведенным коэффициентом направленности* антенны и описывающая угловое распределение излучаемой мощности  $P_0$  (среднее значение  $G$  по единичной сфере равно единице:  $\frac{1}{4\pi} \oint G d\omega = 1$ ). Таким образом,

$$Q_0 = \sigma_{\text{эфф}} \mathcal{S}_0 = P_0 \frac{G \sigma_{\text{эфф}}}{4\pi R^2} = \frac{R_\Sigma |I_0|^2 G \sigma_{\text{эфф}}}{8\pi R^2}.$$

Согласно (20.13) средний квадрат эквивалентной тепловой э. д. с., создаваемой рассматриваемым телом в сечении  $s$  антенны, равен (на  $\omega > 0$ )

$$\hat{\sigma}_\omega^2(s) = 2 \langle |\hat{\phi}(s)|^2 \rangle = \frac{\Theta R_\Sigma(s) G \sigma_{\text{эфф}}}{2\pi^2 R^2}. \quad (20.17)$$

Здесь предполагается, что тело обладает всюду одинаковой температурой ( $\Theta = \text{const}$ ). Если в поле излучения антенны находится несколько тел с различными температурами и на разных удалениях от антенны, то

$$\hat{\sigma}_\omega^2(s) = \frac{R_\Sigma(s)}{2\pi^2} \sum_\alpha \frac{\Theta_\alpha G_\alpha \sigma_{\text{эфф}\alpha}}{R_\alpha^2}. \quad (20.18)$$

Для абсолютно черных тел эффективное сечение совпадает с геометрическим ( $\sigma_{\text{эфф}} = \sigma$ , предполагается, что линейные размеры  $\sigma$  гораздо больше длины волны  $\lambda$ ) и, следовательно,  $\sigma_{\text{эфф}}/R^2 = \sigma/R^2 = 0$ , где  $0$  — телесный угол, под которым тело видно из места расположения антенны. Формула (20.18) запишется тогда

в виде

$$\hat{\mathcal{E}}_{\omega}^a(s) = \frac{2R_{\Sigma}(s)}{\pi} \sum_{\alpha} \Theta_{\alpha} G_{\alpha} \frac{o_{\alpha}}{4\pi}$$

или даже

$$\hat{\mathcal{E}}_{\omega}^a(s) = \frac{2R_{\Sigma}(s)}{\pi} \int \Theta G \frac{do}{4\pi}$$

(при условии, что  $do \gg \lambda^2/R^2$ ). Если температура всех тел одинакова, то мы получаем

$$\hat{\mathcal{E}}_{\omega}^a(s) = \frac{2}{\pi} \Theta R_{\Sigma}(s) \frac{1}{4\pi} \int G do.$$

При сплошном заполнении периферии черными телами со всюду одинаковой температурой (черная оболочка) мы возвращаемся к формуле (20.16).

## § 21. Равновесное тепловое поле. Равновесная форма ФДТ

До сих пор мы рассматривали главным образом *неравновесные* задачи — о поле, создаваемом нагретыми телами в окружающей среде<sup>1)</sup>, которая считалась либо прозрачной (и поэтому не вносящей никакого вклада в тепловые потери вспомогательного дифракционного поля), либо настолько холодной, что ее собственным флуктуационным полем можно было пренебречь. В этом параграфе мы обратимся к случаю, когда *во всем объеме полного поля температура одинакова*, т. е. будем рассматривать *равновесное* флуктуационное поле. Если, в частности, речь идет о телах, окруженных прозрачной средой, не заключенной в идеальную зеркальную оболочку, то мы будем предполагать, что на достаточно больших расстояниях поле ограничено полностью поглощающей оболочкой, имеющей ту же температуру, что и нагретые тела.

Общие формулы (17.11) — (17.13) для пространственных корреляционных функций спектральных амплитуд флуктуационного поля, создаваемого равномерно нагретыми телами, можно объединить в одну формулу, если обозначить через  $A$  и  $B$  какие-либо две из шести компонент  $E$  и  $H$ . Тогда вместо (17.11) — (17.13) можно написать

$$\langle A(r_1) B^*(r_2) \rangle = \pm \frac{2}{\pi} \Theta(\omega, T) Q_{AB}, \quad (21.1)$$

<sup>1)</sup> Напомним, что состояние самих нагретых тел необходимо предполагать при этом квазиравновесным.

Верхний знак отвечает случаю обеих электрических или обеих магнитных компонент, нижний — одной электрической и другой магнитной;  $Q_{0AB}$  — смешанные потери дифракционного поля, создаваемого единичными точечными источниками соответствующего типа, помещенными в точках  $r_1$  и  $r_2$ .

Кирхгофская форма ФДТ (21.1) справедлива, конечно, и для равновесного поля, но под  $Q_{0AB}$  надо понимать теперь смешанные потери дифракционного поля во всем занимаемом им пространстве. Это обстоятельство позволяет существенно упростить выражение для  $Q_{0AB}$ , воспользовавшись так называемой *комплексной леммой Лоренца*. Эта лемма, в известном смысле аналогичная теореме взаимности, отличается от последней тем, что она связывает «накрест» поле одной системы источников с комплексно сопряженным полем другой системы источников.

Написав, как и при выводе теоремы взаимности (17.2), две системы уравнений Максвелла (16.1) — одну для поля и источников 1, другую для поля и источников 2, — умножим уравнения первой системы соответственно на  $-H_2^*$  и  $E_2^*$ , а комплексно сопряженные уравнения второй системы на  $-H_1$  и  $E_1$  и сложим результаты. Это приводит к равенству

$$\begin{aligned} -\frac{c}{4\pi} \operatorname{div} \{[E_1 H_2^*] + [E_2^* H_1]\} = \\ = j_{e1} E_2^* + j_{m1} H_2^* + j_{e2}^* E_1 + j_{m2}^* H_1 - \\ - \frac{i\omega}{4\pi} (D_1 E_2^* - E_1 D_2^* + B_1 H_2^* - H_1 B_2^*). \end{aligned}$$

При интегрировании этого равенства по всему объему  $V$  полного поля интеграл от  $\operatorname{div}$  обращается в нуль и мы получаем лемму Лоренца:

$$\begin{aligned} -\frac{i\omega}{16\pi} \int_V (D_1 E_2^* - E_1 D_2^* + B_1 H_2^* - H_1 B_2^*) d^3r = \\ = -\frac{1}{4} \int_V (j_{e1} E_2^* + j_{m1} H_2^* + j_{e2}^* E_1 + j_{m2}^* H_1) d^3r. \quad (21.2) \end{aligned}$$

Выражение, стоящее в левой части (21.2), если учесть, что  $D_{1\alpha} = \epsilon_{\alpha\beta} E_{1\beta}$ ,  $D_{2\beta}^* = \epsilon_{\beta\alpha}^* E_{2\alpha}^*$  (и аналогично для  $B_1$  и  $B_2^*$ ), представляет собой не что иное, как смешанные тепловые потери  $Q_{12}$  полей 1 и 2 во всем пространстве (см. (17.9), (17.10)), и, значит,

$$Q_{12} = -1/4 \int_V (j_{e1} E_2^* + j_{m1} H_2^* + j_{e2}^* E_1 + j_{m2}^* H_1) d^3r. \quad (21.3)$$

Этим результатом леммы Лоренца мы теперь и воспользуемся применительно к дифракционным полям, создаваемым единичными точечными источниками.

Возьмем случай, когда  $A$  и  $B$  в (21.1) — компоненты электрической напряженности теплового поля  $E$ . Тогда

$$j_1 \equiv j_{e1} = I_1 \delta(r - r_1), \quad j_2 \equiv j_{e2} = I_2 \delta(r - r_2), \quad j_{m1} = j_{m2} = 0$$

и (21.3) дает

$$Q_{ee}(r_1, I_1; r_2, I_2) = -1/4 \{I_1 E_{02}^*(r_1) + I_2 E_{01}(r_2)\}. \quad (21.4)$$

Но из обычной теоремы взаимности (17.2) следует, что в этом же случае

$$I_1 E_{02}(r_1) = I_2 E_{01}(r_2).$$

Исключая из (21.4) при помощи последнего равенства либо  $E_{01}$ , либо  $E_{02}^*$ , можно записать потери  $Q_{ee}$  в следующих двух формах:

$$Q_{ee}(r_1, I_1; r_2, I_2) = -1/2 \operatorname{Re} \{I_1 E_{02}(r_1)\} = -1/2 \operatorname{Re} \{I_2 E_{01}(r_2)\}.$$

Подстановка этих выражений в (17.12) дает

$$\begin{aligned} \langle E_{I_1}(r_1) E_{I_2}^*(r_2) \rangle &= -\frac{1}{\pi} \Theta(\omega, T) \operatorname{Re} \{I_1 E_{02}(r_1)\} = \\ &= -\frac{1}{\pi} \Theta(\omega, T) \operatorname{Re} \{I_2 E_{01}(r_2)\}. \end{aligned} \quad (21.5)$$

Применение (21.3) к случаю двух компонент магнитной напряженности теплового поля  $H$  приводит к аналогичному выражению для  $Q_{mm}$  и, в соответствии с (17.13), к результату:

$$\begin{aligned} \langle H_{I_1}(r_1) H_{I_2}^*(r_2) \rangle &= -\frac{1}{\pi} \Theta(\omega, T) \operatorname{Re} \{I_1 H_{02}(r_1)\} = \\ &= -\frac{1}{\pi} \Theta(\omega, T) \operatorname{Re} \{I_2 H_{01}(r_2)\}. \end{aligned} \quad (21.6)$$

Наконец, в случае, когда  $A$  — компонента  $E$ , а  $B$  — компонента  $H$ , мы получаем  $Q_{em}$  и из (17.11) следует, что

$$\begin{aligned} \langle E_{I_1}(r_1) H_{I_2}^*(r_2) \rangle &= -\frac{i}{\pi} \Theta(\omega, T) \operatorname{Im} \{I_1 E_{02}(r_1)\} = \\ &= \frac{i}{\pi} \Theta(\omega, T) \operatorname{Im} \{I_2 H_{01}(r_2)\}, \end{aligned} \quad (21.7)$$

т. е. корреляционные функции между электрической и магнитной напряженностями чисто мнимые.

Формулы (21.5) — (21.7) показывают, что с точностью до множителя  $-\Theta/\pi$  корреляционные функции равновесного поля совпадают с вещественными или мнимыми частями соответствующих функций Грина. Эти формулы можно назвать равновесной формой ФДТ. Она избавляет от необходимости вычислять после нахождения вспомогательного дифракционного поля его тепловые потери. Достаточно взять вещественную или мнимую часть самих напряженностей этого поля. Вместе с тем линейная связь между вторыми моментами теплового поля и напряженностями дифрак-

дионного поля сразу же позволяет использовать аппарат теории аналитических функций комплексного переменного при нахождении тех или иных интегральных (по спектру) эффектов в равновесном поле, т. е. при интегрировании корреляционных функций теплового поля по частоте  $\omega$ .

Располагая функциями корреляции спектральных компонент флуктуационного поля и зная, в частности, средние значения произведений этих компонент в одной точке, мы имеем возможность находить не только средние плотности энергии и ее потока, но и другие средние билинейные величины. Мы можем, например, вычислить средние значения компонент максвелловского тензора натяжений, т. е. найти механические (пондеромоторные) силы, с которыми равновесное поле действует на тела. Сказанное относится как к квазиравновесным полям, когда в нашем распоряжении обобщенный закон Кирхгофа (21.1), так и к равновесным, когда можно пользоваться равновесной формой ФДТ (21.5) и (21.6). Однако эти интересные физические применения указанных теорем [6] не имеют прямого отношения к радиофизике, и поэтому мы ограничимся лишь немногими замечаниями.

На первый взгляд может показаться странным, что даже в *равновесном* случае флуктуационное электромагнитное поле обуславливает действие сил на тела (в том числе и на идеальные проводники). Но ничего удивительного здесь нет просто потому, что, говоря о равновесии, мы постоянно имели в виду только *тепловое* равновесие, т. е. одинаковую во всей рассматриваемой системе температуру  $T$  (в частности,  $T=0$ ). Это вовсе не исключает действия электродинамических сил, которые либо уравновешены внешними связями, либо вызывают перемещение тел к *механически равновесной* конфигурации. Поэтому нет ничего удивительного и в том, например, что равновесное поле в полупространстве над идеальным плоским зеркалом оказывает на это зеркало давление, спектральная плотность которого в случае недиспергирующей среды равна

$$p_{\omega} = -\frac{1}{4\pi} \{ \epsilon \langle |E_n|^2 \rangle - \mu \langle |H_t|^2 \rangle \} = \frac{\Theta k^2}{3\pi^2 \omega} = \frac{u_m}{3}.$$

Как всегда, член  $\hbar\omega/2$  в  $\Theta(\omega, T)$  (т. е. нулевые колебания) дает при интегрировании по  $\omega$  расходящееся выражение. Однако эта часть давления компенсируется «нулевым» давлением поля по другую сторону от зеркала, где равновесное излучение может, вообще говоря, иметь другую температуру. Результирующая сила, действующая на зеркало, определяется разностью давлений по обе стороны и не содержит в данном примере вклада нулевых колебаний.

Это не означает, что нулевые колебания вообще не создают пондеромоторных сил. В случае плоского зеркала геометричес-

кие условия симметричны, одинаковы по обе стороны от зеркала. Если же взять, скажем, образованный двумя идеальными зеркалами двугранный угол (рис. 19), то структура «нулевых» стоячих волн внутри угла и вне его будет различна. В результате получается, что даже при  $T=0$  на зеркала действует «схлопывающий» интегральный (по  $\omega$ ) момент. Интегрирование по  $\omega$  надо при этом распространять не до  $\omega=\infty$ , а лишь до некоторой



Рис. 19.

частоты  $\omega_m \ll \sigma$  ( $\sigma$  — проводимость; неравенство является условием того, что при нормальном скин-эффекте металл еще можно считать идеально проводящим).

Разумеется, флуктуационное электромагнитное поле порождает силы, действующие и на поглощающие тела. Более того, поскольку поверхности таких тел «выстланы» слоем ближнего (квазистационарного) теплового поля, сильно нарастающего при приближении к поверхности, эффект механического взаимодействия (сил сцепления) становится особенно большим в случае малых зазоров между поверхностями тел. На такого рода явления обратил внимание Е. М. Лифшиц, исследовавший их в работе [7] (см. также [11], [12] и § 18 в [6]).

Теория макроскопических сил сцепления строилась ранее на основе элементарного закона ван-дер-ваальсовых сил попарного взаимодействия между атомами или молекулами, что заранее ограничивало результат случаев разреженных сред. Чисто феноменологическая теория, основанная на флуктуационной электродинамике, снимает это ограничение. Напротив, исходя из выражения для макроскопической силы сцепления, можно в случае разреженных сред сделать обратное заключение — о законе попарного взаимодействия отдельных нейтральных атомов и молекул. Такой путь, как это ни парадоксально на первый взгляд, оказывается проще, чем прямой квантовомеханический расчет для двух нейтральных частиц, при котором закон взаимодействия получается лишь в высоких порядках при вычислениях методом возмущений.

## § 22. Тепловое поле в гиротропных телах

Многие среды (ферриты, плазма), находясь в достаточно сильном постоянном (во времени) внешнем магнитном поле  $B_0$ , становятся *гиротропными*<sup>1)</sup>. Это означает, в частности, что в одно-

<sup>1)</sup> Под «достаточно сильным» понимается поле, модуль напряженности которого  $B_0$  существенно больше стандартных напряженностей флуктуационного поля. Поэтому при линеаризации уравнений Максвелла относительно этих флуктуационных полей сохраняет смысл учет зависимости параметров среды от  $B_0$ .



родной среде и однородном поле  $B_0$  собственными волнами являются монохроматические волны с *круговой* (правой и левой) поляризацией, т. е. только эти волны распространяются без изменения вида поляризации. Если же в среду введена, скажем, линейно поляризованная волна, то по мере распространения ее плоскость поляризации будет, вообще говоря, поворачиваться вокруг направления распространения — за счет разницы фазовых скоростей волн с правой и левой круговой поляризацией, на которые можно разложить линейно поляризованную волну.

В то время как в анизотропных средах тензоры проницаемостей симметричны:

$$\begin{aligned}\epsilon_{\alpha\beta}(\omega, \mathbf{r}) &= \epsilon_{\beta\alpha}(\omega, \mathbf{r}), \\ \mu_{\alpha\beta}(\omega, \mathbf{r}) &= \mu_{\beta\alpha}(\omega, \mathbf{r}),\end{aligned}$$

в гиротропной среде, будучи функциями  $B_0$ , они обладают симметрией с *переменой знака*  $B_0$ :

$$\begin{aligned}\epsilon_{\alpha\beta}(\omega, \mathbf{r}, B_0) &= \epsilon_{\beta\alpha}(\omega, \mathbf{r}, -B_0), \\ \mu_{\alpha\beta}(\omega, \mathbf{r}, B_0) &= \mu_{\beta\alpha}(\omega, \mathbf{r}, -B_0).\end{aligned}$$

Тем самым, для гиротропных сред обычная теорема взаимности (17.2) несправедлива. Ее заменяет равенство, связывающее «накрест» источники и напряженности поля 1 в среде, находящейся в магнитном поле  $B_0(j_{e1}, j_{m1}, E_1, H_1)$ , и источники и напряженности поля 2 в той же среде, но с *обращенным внешним магнитным полем* —  $B_0$ . Будем отмечать электромагнитные величины, относящиеся к полю 2 в такой «обращенной среде», значком «тильда» ( $\tilde{j}_{e2}, \tilde{j}_{m2}, \tilde{E}_2, \tilde{H}_2$ ). Вместо (17.2) для гиротропной среды выполняется теорема

$$\int_V (\tilde{E}_2 j_{e1} - \tilde{H}_2 j_{m1}) d^3r = \int_V (E_1 \tilde{j}_{e2} - H_1 \tilde{j}_{m2}) d^3r. \quad (22.1)$$

Формулы (16.4) для корреляционных функций *сторонних токов* справедливы, конечно, и в гиротропной среде, но обобщенный закон Кирхгофа (21.1), дающий корреляцию *напряженностей флуктуационного поля*, уже теряет силу, так как при выводе этого закона была использована теорема взаимности (17.2).

Если повторить весь вывод, но опираясь на равенство (22.1), то мы придем к следующему результату, обобщающему формулы (21.1) на случай гиротропных сред:

$$\langle A(\mathbf{r}_1) B^*(\mathbf{r}_2) \rangle = \pm \frac{2}{\pi} \Theta(\omega, T) \bar{Q}_{\alpha\lambda B^*}, \quad (22.2)$$

а при достаточном плавном пространственном изменении температуры —

$$\langle A(r_1) B^*(r_2) \rangle = \pm \frac{2}{\pi} \int \Theta(\omega, T, r) d\tilde{Q}_{AB}. \quad (22.3)$$

Таким образом, корреляция спектральных амплитуд флуктуационного поля в гиротропной среде, находящейся во внешнем магнитном поле  $B_0(r)$ , определяется тепловыми потерями вспомогательного дифракционного поля в обращенной среде, т. е. *потери надо вычислять при изменении знака внешнего магнитного поля ( $B_0(r) \rightarrow -B_0(r)$ )*<sup>1)</sup>. Формулы (22.2), (22.3) универсальны в том смысле, что они не налагают никаких ограничений ни на размеры и форму тел, ни на их электрические и магнитные свойства, включая и гиротропию, зависящую также от интенсивности и конфигурации постоянного (во времени) подмагничивающего поля  $B_0$ .

Формулы (22.2), (22.3) облегчают решение флуктуационных задач для гиротропных тел в не меньшей степени, чем формулы (21.1) в отсутствие гиротропии. Разумеется, необходимое для вычисления тепловых потерь решение вспомогательной дифракционной задачи связано, вообще говоря, с большими трудностями, чем для негиротропных тел и сред. Достаточно указать, что даже нахождение френелевских коэффициентов отражения плоской волны от однородного гиротропного полупространства представляет собой при произвольной ориентации однородного же поля  $B_0$  чрезвычайно громоздкую задачу. Обычно здесь ограничиваются поэтому лишь частными случаями (поле  $B_0$  перпендикулярно или параллельно границе, первичная волна падает нормально). Круг точно решаемых дифракционных задач, т. е. допускающих нахождение точных функций Грина, ограничен для гиротропных тел еще более жестко, чем для тел изотропных, но это лежит в природе вещей, а не в методе решения флуктуационных задач. Нахождение статистических характеристик теплового поля в принципе столь же просто.

Естественно, что все формулы, выводимые на основе обобщенного закона Кирхгофа, претерпевают при рассмотрении гиротропных тел такое же изменение, какое содержится в (22.2). Так, например, для потока энергии теплового излучения в телесный угол  $d\omega$  в волновой зоне *гиротропного* тела вместо формулы (18.15) будет теперь

$$dP_{\omega} = \frac{\Theta(\omega, T)}{\pi \lambda^2} \tilde{\sigma}_{\omega\phi\phi} d\omega, \quad (22.4)$$

<sup>1)</sup> Не только обобщенный, но и классический закон Кирхгофа для гиротропных сред долгое время оставался неизвестным. Закон (22.3) был получен М. Л. Левиным лишь в конце 50-х годов, а опубликован еще позднее, в книге [6].

где  $\tilde{\sigma}_{\perp\perp}$  — эффективный поперечник поглощения данного тела при *обращенном* поле  $-\mathbf{B}_0$ . Волноводная форма закона Кирхгофа (19.5) тоже заменится на

$$P_{\omega n}^{\pm} = \frac{\Theta}{2\pi} \tilde{A}_n^{\mp}. \quad (22.5)$$

Ряд примеров применения формул (22.2) — (22.5) см. в книге [6], §§ 21—23.

Что касается равновесной формы ФДТ (§ 21), то для нее при наличии гиротропия дело обстоит следующим образом. Формулы (21.5) — (21.7) были получены при помощи обобщенного закона Кирхгофа (21.1) и комплексной леммы Лоренца, т. е. результата (21.4). Лемма Лоренца справедлива для *любых* сред с *линейными* материальными уравнениями, но (21.1) заменяется в случае гиротропных сред на (22.2). Поэтому, например, вместо формулы (см. (21.5))

$$\langle E_i(r_1) E_i^*(r_2) \rangle = -\frac{1}{2\pi} \Theta(\omega, T) \{I_1 E_{0i}^*(r_1) + I_2 E_{0i}(r_2)\}$$

мы получим теперь

$$\langle E_i(r_1) E_i^*(r_2) \rangle = -\frac{1}{2\pi} \Theta(\omega, T) \{I_1 \tilde{E}_{0i}^*(r_1) + I_2 \tilde{E}_{0i}(r_2)\}, \quad (22.6)$$

где справа входят функции Грина в обращенной среде.

Но из теоремы (22.1), т. е. из обобщенной на гиротропные среды теоремы взаимности, следует, что при рассматриваемых точечных источниках

$$I_1 \tilde{E}_{0i}(r_1) = I_2 E_{0i}(r_2). \quad (22.7)$$

При помощи этого равенства можно исключить из (22.6), скажем  $\tilde{E}_{0i}(r_1)$ , что дает

$$\langle E_i(r_1) E_i^*(r_2) \rangle = -\frac{1}{2\pi} \Theta(\omega, T) I_2 \{E_{0i}^*(r_1) + \tilde{E}_{0i}(r_2)\}. \quad (22.8)$$

Это смешанная форма корреляционной функции, содержащая функции Грина как в исходной среде, так и в среде с обращенным подмагничивающим полем. При  $\mathbf{B}_0 = 0$  формулы смешанного типа (в частности, (22.8)) непосредственно переходят в полученные ранее формулы для негиротропных тел (в частности, (21.5)).

Можно также, пользуясь (22.7), заменить в (22.6) обе напряженности с тильдами на напряженности в исходной среде (без обращения  $\mathbf{B}_0$ ), и тогда

$$\langle E_i(r_1) E_i^*(r_2) \rangle = -\frac{1}{2\pi} \Theta(\omega, T) \{I_2 E_{0i}^*(r_2) + I_1 E_{0i}(r_1)\}. \quad (22.9)$$

Аналогично формулам (22.8) и (22.9) записываются и функции для компонент  $\mathbf{H}$ , и взаимные функции корреляции между компонентами  $\mathbf{E}$  и  $\mathbf{H}$ .

Таким образом, в случае равновесного поля гиротропность не вносит никаких принципиальных усложнений. Как и для негиротропных тел, надо знать только функции Грина, хотя их фактическое вычисление, конечно, значительно сложнее, чем для изотропных тел.

Ясно, что наличие гиротропности не должно нарушать универсальную связь между *интегральным* излучением и поглощением одного и того же тела, вытекающую из второго начала термодинамики. Если, например, тело окружено достаточно удаленной абсолютно черной оболочкой с той же температурой, что и у тела, то тепловое равновесие между телом и оболочкой должно иметь место независимо от того, обладает тело гиротропией или нет. Это значит, что интегральная (по телесному углу) суммарная (по обеим поляризациям) интенси́вность излучения тела должна быть в обоих случаях одинакова ( $P_{\omega} = \bar{P}_{\omega}$ ). Но в отсутствие гиротропии, согласно (18.15),

$$P_{\omega} = \frac{\Theta}{\pi\lambda^2} \oint \sigma_{\omega\phi\phi} d\omega,$$

а при ее наличии, согласно (22.4),

$$\bar{P}_{\omega} = \frac{\Theta}{\pi\lambda^2} \oint \bar{\sigma}_{\omega\phi\phi} d\omega.$$

Из равенства  $P_{\omega} = \bar{P}_{\omega}$  следует, что

$$\oint \sigma_{\omega\phi\phi} d\omega = \oint \bar{\sigma}_{\omega\phi\phi} d\omega. \quad (22.10)$$

Точно так же суммарное излучение тела, находящегося в волноводе между двумя черными пробками, вправо и влево

$$P_{\omega} = P_{\omega}^{+} + P_{\omega}^{-} = \sum_n (P_{\omega n}^{+} + P_{\omega n}^{-})$$

не должно зависеть от наличия гиротропии, т. е. должно быть  $P_{\omega} = \bar{P}_{\omega}$ . Суммируя выражения (19.4) по всем номерам  $n$  бегущих волн, получаем

$$P_{\omega}^{+} = \frac{\Theta}{2\pi} \left\{ \sum_n A_n^{-} + \sum_{m,n} (D_{nm}^{-} - D_{mn}^{+}) \right\},$$

$$P_{\omega}^{-} = \frac{\Theta}{2\pi} \left\{ \sum_n A_n^{+} + \sum_{m,n} (D_{nm}^{+} - D_{mn}^{-}) \right\}.$$

Конечно, в двойных суммах по  $m$  и  $n$  эти индексы можно переставлять. Поэтому

$$P_{\omega} = P_{\omega}^{+} + P_{\omega}^{-} = \frac{\Theta}{2\pi} \sum_n (A_n^{-} + A_n^{+}).$$

Из равенства  $P_{\omega} = \bar{P}_{\omega}$  вытекает соотношение, аналогичное (22.10):

$$\sum_n (A_n^{-} + A_n^{+}) = \sum_n (\bar{A}_n^{-} + \bar{A}_n^{+}). \quad (22.11)$$

### § 23. Тепловое поле в среде с пространственной дисперсией

До сих пор мы ограничивались материальными уравнениями (14.2), описывающими среду (вообще говоря, неоднородную) и отсутствие пространственной дисперсии: индукции  $\mathbf{D}$  и  $\mathbf{B}$  в точке  $\mathbf{r}$  зависели от свойств и предыстории как среды (от  $\epsilon$  и  $\mu$ ), так и поля (напряженностей  $\mathbf{E}$  и  $\mathbf{H}$ ) только в той же точке  $\mathbf{r}$ . Если же состояние среды в точке  $\mathbf{r}$  зависит от предыстории (в чем заключается *временная* нелокальность) не только в этой же точке, но и в некоторой ее окрестности (*пространственная* келокальность), то материальные уравнения надо брать более общего вида:

$$\begin{aligned} \mathbf{D}(\mathbf{r}, t) &= \int_{-\infty}^{+\infty} dt' \int_{-\infty}^{+\infty} \epsilon(t-t', \mathbf{r}, \mathbf{r}') \mathbf{E}(t', \mathbf{r}') d^3r', \\ \mathbf{B}(\mathbf{r}, t) &= \int_{-\infty}^{+\infty} dt' \int_{-\infty}^{+\infty} \mu(t-t', \mathbf{r}, \mathbf{r}') \mathbf{H}(t', \mathbf{r}') d^3r'. \end{aligned} \quad (23.1)$$

Зависимость ядер  $\epsilon$  и  $\mu$  от разности  $t-t'$  отражает *стационарность* интересующих нас процессов, т. е. предполагает неизменность свойств среды во времени (если бы среда менялась во времени, то  $\epsilon$  и  $\mu$  зависели бы от  $t$  и  $t'$  порознь). Зависимость  $\epsilon$  и  $\mu$  от  $\mathbf{r}$  и  $\mathbf{r}'$  отвечает пространственно *неоднородной* среде. В однородной среде ядра  $\epsilon$  и  $\mu$  зависят только от разности  $\mathbf{r}=\mathbf{r}-\mathbf{r}'$ . Если область пространственной нелокальности стягивается в точку, то

$$\epsilon(t-t', \mathbf{r}, \mathbf{r}') \rightarrow \epsilon(t-t', \mathbf{r}) \delta(\mathbf{r}-\mathbf{r}'),$$

и аналогично для  $\mu$ . Материальные уравнения (23.1) тотчас же переходят при этом в локальные уравнения (14.2)<sup>1)</sup>.

В (23.1), как и в (14.2), имеются в виду изотропная и негиротропная среды. Для анизотропной или гиротропной среды соотношение между  $\mathbf{D}$  и  $\mathbf{E}$  устанавливается тензором диэлектри-

<sup>1)</sup> Состояние среды в момент  $t$  в точке  $\mathbf{r}$  может зависеть только от событий, которые находятся внутри обращенного в прошлое светового конуса с вершиной в «точке»  $(t, \mathbf{r})$ . Иначе говоря, это события, которые предшествуют моменту  $t$  по времени и создают поля, распространяющиеся не быстрее света в вакууме (принцип причинности). Таким образом, интегрирование по  $t'$  и  $\mathbf{r}'$  в (23.1) и в последующей формуле (23.2a) в бесконечных пределах предполагает, что функции  $\epsilon$ ,  $\mu$  и  $\epsilon_{\alpha\beta}$  обращаются вне указанного конуса в нуль. Это налагает определенные ограничения и на их фж-амплитуды, а именно устанавливает определенную связь между вещественными и мнимыми частями этих функций. В отсутствие пространственной дисперсии такая связь выражается формулами Крамерса—Кронига ([3], § 62), а при ее наличии—более общими формулами [13].

ческой проницаемости  $\epsilon_{\alpha\beta}$  ( $\alpha, \beta = 1, 2, 3$ ), так что первое уравнение (23.1) заменяется на

$$D_{\alpha}(t, \mathbf{r}) = \int_{-\infty}^{+\infty} dt' \int_{-\infty}^{+\infty} \epsilon_{\alpha\beta}(t-t', \mathbf{r}, \mathbf{r}') E_{\beta}(t', \mathbf{r}') d^3\mathbf{r}'. \quad (23.2a)$$

Что касается второго уравнения (23.1), то из общих соображений следует, что введение еще одного тензора  $\mu_{\alpha\beta}$  было бы излишним. Дело в том, что при введении  $\mathbf{D} = \mathbf{E} + 4\pi\mathbf{P}$  и  $\mathbf{B} = \mathbf{H} + 4\pi\mathbf{M}$  мы описываем состояние среды поляризацией  $\mathbf{P}$  и намагничиванием  $\mathbf{M}$ , разбивая для этого индуцированные в среде токи и заряды на «сорты» (связанные и свободные заряды с соответствующими поляризационным током и током проводимости, а кроме того, молекулярные амперовы токи). Но в нелокальной среде проще и естественнее включить в  $\mathbf{D}$  все виды наведенного тока, так что

$$\mathbf{B} = \mathbf{H}, \quad (23.26)$$

а для  $\mathbf{D}$  справедливо материальное уравнение (23.2a). Конечно, определения  $\mathbf{D}$  и  $\mathbf{B}$  (т. е. их связь с микрополями и микротоками) при этом *иные*, чем в (23.1), так что при отсутствии пространственной дисперсии уравнения (23.2a, б) *не переходят* в (23.1).

Возможность полного описания среды материальными уравнениями (23.2a, б) показывает, что достаточно *одного* тензора  $\epsilon_{\alpha\beta}$  при  $\mu_{\alpha\beta} = \delta_{\alpha\beta}$  (подробней см. [8, 14]). Заметим, что в *толще* нелокальной среды, даже если она однородна и изотропна, тензор  $\epsilon_{\alpha\beta}(t-t', \mathbf{r}, \mathbf{r}')$  в (23.2a) не вырождается в диагональный, а имеет общий для данного случая вид:

$$\epsilon_{\alpha\beta}(\tau, \rho) = \left( \delta_{\alpha\beta} - \frac{\rho_{\alpha}\rho_{\beta}}{\rho^2} \right) \epsilon_t(\tau, \rho) + \frac{\rho_{\alpha}\rho_{\beta}}{\rho^2} \epsilon_l(\tau, \rho), \quad (23.3)$$

где  $\tau = t-t'$ ,  $\rho = \mathbf{r}-\mathbf{r}'$ , т. е. содержит *две* независимые скалярные функции  $\epsilon_t$  и  $\epsilon_l$  — поперечную и продольную (по отношению к  $\rho$ ) диэлектрические проницаемости. Отсюда ясно, что для *однородной* и *изотропной* нелокальной среды допустимы материальные уравнения и вида (23.1), поскольку в них тоже фигурируют только две независимые скалярные функции  $\epsilon(\tau, \rho)$  и  $\mu(\tau, \rho)$ . О связи между  $\epsilon_t$ ,  $\epsilon_l$  и  $\epsilon$ ,  $\mu$  мы скажем несколько дальше.

Выше мы не случайно подчеркивали, что выражение (23.3) для  $\epsilon_{\alpha\beta}$  относится к толще среды, т. е. к областям, достаточно удаленным от ее границ. При приближении к границе среды  $\epsilon_{\alpha\beta}$  не может оставаться функцией только разности  $\rho = \mathbf{r}-\mathbf{r}'$ , так как направление по нормали к границе становится выделенным. Иначе говоря, прилегающий к границе слой с толщи-

ной порядка радиуса нелокальности неоднороден и анизотропен по своим электродинамическим свойствам. Это в особой степени усложняет электродинамическую часть задачи, т. е. нахождение вспомогательных дифракционных полей (функций Грина) или же плотности их джоулевых потерь, если необходимо учитывать и указанный приграничный слой. Что же касается собственно флуктуационной части задачи, то в ней все обстоит так же, как и в случае локальных сред, поскольку и обобщенный закон Кирхгофа (17.11) — (17.13), и равновесная форма ФДТ (21.5) — (21.7) справедливы при *любой* форме материальных уравнений.

Мы ограничимся далее простейшим случаем *однородной* и *изотропной* нелокальной среды, причем среды *неограниченной*, когда во всем пространстве можно пользоваться как  $\epsilon_{\alpha\beta}(\tau, \rho)$  в форме (23.3), так и скалярами  $\epsilon(\tau, \rho)$ ,  $\mu(\tau, \rho)$ . Целесообразно при этом перейти к трансформантам Фурье как по  $\tau$ , так и по  $\rho$ :

$$\epsilon(\tau, \rho) = \frac{1}{(2\pi)^4} \iint_{-\infty}^{+\infty} \epsilon(\omega, \mathbf{k}) \exp[-i(\omega\tau - \mathbf{k}\rho)] d\omega d^3\mathbf{k}, \quad (23.4)$$

и аналогично для  $\mu(\tau, \rho)$ . Такое же преобразование Фурье для тензора  $\epsilon_{\alpha\beta}(\tau, \rho)$  содержит под интегралом трансформанту

$$\epsilon_{\alpha\beta}(\omega, \mathbf{k}) = \left( \delta_{\alpha\beta} - \frac{\kappa_\alpha \kappa_\beta}{\kappa^2} \right) \epsilon_t(\omega, \mathbf{k}) + \frac{\kappa_\alpha \kappa_\beta}{\kappa^2} \epsilon_l(\omega, \mathbf{k}), \quad (23.5)$$

где  $\epsilon_t(\omega, \mathbf{k})$  и  $\epsilon_l(\omega, \mathbf{k})$  — поперечная и продольная (по отношению к  $\mathbf{k}$ ) проицируемости, конечно, *не являющиеся* преобразованиями Фурье от  $\epsilon_t(\tau, \rho)$  и  $\epsilon_l(\tau, \rho)$  в (23.3). Связь скалярных функций  $\epsilon(\omega, \mathbf{k})$  и  $\mu(\omega, \mathbf{k})$  с функциями  $\epsilon_t(\omega, \mathbf{k})$  и  $\epsilon_l(\omega, \mathbf{k})$  очень проста [8]:

$$\epsilon(\omega, \mathbf{k}) = \epsilon_t(\omega, \mathbf{k}), \quad (23.6)$$

$$\frac{1}{\mu(\omega, \mathbf{k})} = \frac{k_0^2}{\kappa^2} [\epsilon_t(\omega, \mathbf{k}) - \epsilon_l(\omega, \mathbf{k})], \quad k_0 = \frac{\omega}{c}.$$

Вторая из этих формул показывает, что  $\mu(\omega, \mathbf{k})$  не может обращаться в нуль.

Целесообразность использования при указанных условиях пространственно-временных трансформант Фурье ( $\omega\mathbf{k}$ -амплитуд) или пространственных трансформант Фурье от спектральных амплитуд, вытекает из того, что при этом все линейные уравнения для полей превращаются в *алгебраические* линейные уравнения для их  $\omega\mathbf{k}$ -амплитуд. В частности, уравнения Максвелла (14.6), если мы полагаем

$$\mathbf{E}(t, \mathbf{r}) = \iint_{-\infty}^{+\infty} \mathbf{E}(\omega, \mathbf{k}) \exp[-i(\omega t - \mathbf{k}\mathbf{r})] d\omega d^3\mathbf{k} \quad (23.7)$$

и аналогично для  $\mathbf{H}$ , индукций  $\mathbf{D}$ ,  $\mathbf{B}$  и сторонних токов  $\mathbf{J}_0$ ,  $\mathbf{J}_m$ , принимают для  $\omega$ -амплитуд следующий вид (аргументы  $\omega$  и  $\mathbf{k}$  мы для краткости опускаем):

$$\begin{aligned} i[\mathbf{kH}] &= -ik_0\mathbf{D} + \frac{4\pi}{c}\mathbf{J}_e, \\ i[\mathbf{kE}] &= ik_0\mathbf{B} - \frac{4\pi}{c}\mathbf{J}_m, \end{aligned} \quad (23.8)$$

а материальные уравнения для  $\omega$ -амплитуд, полученные из (23.1) с учетом (23.4) и (23.7) (среда однородная), сводятся к соотношениям

$$\mathbf{D}(\omega, \mathbf{k}) = \varepsilon(\omega, \mathbf{k}) \mathbf{E}(\omega, \mathbf{k}), \quad \mathbf{B}(\omega, \mathbf{k}) = \mu(\omega, \mathbf{k}) \mathbf{H}(\omega, \mathbf{k}). \quad (23.9)$$

Наличие пространственной дисперсии отражено в том, что  $\varepsilon$  и  $\mu$  зависят от  $\mathbf{k}$  (для изотропной среды они зависят только от  $k = |\mathbf{k}|$ ).

Алгебраическую форму принимает и ФДТ. В общем случае многомерного флуктуационного поля  $\xi^{(j)}(t, \mathbf{r})$  и сопряженного по Лагранжу силового поля  $f^{(j)}(t, \mathbf{r})$  операторные уравнения (15.12) переходят для спектральных амплитуд в результате пространственного преобразования Фурье в алгебраические уравнения

$$\begin{aligned} \xi^{(j)}(\omega, \mathbf{k}) &= \sum_{\mathbf{k}'} A_{j\mathbf{k}}(-i\omega, i\mathbf{k}) f^{(k)}(\omega, \mathbf{k}'), \\ f^{(j)}(\omega, \mathbf{k}) &= \sum_{\mathbf{k}'} A_{j\mathbf{k}}^{-1}(-i\omega, i\mathbf{k}) \xi^{(k)}(\omega, \mathbf{k}'), \end{aligned} \quad (23.10)$$

а вытекающие из ФДТ выражения для корреляционных матриц  $\omega$ -амплитуд  $\xi^{(j)}$  и  $f^{(j)}$  имеют вид (см. задачу 6)

$$\begin{aligned} \langle \xi^{(j)}(\omega, \mathbf{k}) \xi^{(k)*}(\omega, \mathbf{k}') \rangle &\equiv G_{jk}^{(j)}(\omega, \mathbf{k}) \delta(\mathbf{k} - \mathbf{k}') = \\ &= -\frac{i\Theta(\omega, T)}{(2\pi)^4 \omega} \{A_{j\mathbf{k}}(-i\omega, i\mathbf{k}) - A_{k\mathbf{j}}^*(-i\omega, i\mathbf{k})\} \delta(\mathbf{k} - \mathbf{k}'), \end{aligned} \quad (23.11)$$

$$\begin{aligned} \langle f^{(j)}(\omega, \mathbf{k}) f^{(k)*}(\omega, \mathbf{k}') \rangle &\equiv G_{jk}^{(f)}(\omega, \mathbf{k}) \delta(\mathbf{k} - \mathbf{k}') = \\ &= -\frac{i\Theta(\omega, T)}{(2\pi)^4 \omega} \{A_{k\mathbf{j}}^{-1}(-i\omega, i\mathbf{k}) - A_{j\mathbf{k}}^{-1}(-i\omega, i\mathbf{k})\} \delta(\mathbf{k} - \mathbf{k}'), \end{aligned} \quad (23.12)$$

где  $G_{jk}(\omega, \mathbf{k})$  —  $\omega$ -плотности. Преобразование Фурье этих плотностей по  $\mathbf{k}$  дает, конечно, спектральные плотности. Например,

$$\begin{aligned} g_{jk}^{(j)}(\omega, \rho) &\equiv \langle \xi^{(j)}(\omega, \mathbf{r}) \xi^{(k)*}(\omega, \mathbf{r}') \rangle = \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \langle \xi^{(j)}(\omega, \mathbf{k}) \xi^{(k)*}(\omega, \mathbf{k}') \rangle \exp[i(\mathbf{k}\mathbf{r} - \mathbf{k}'\mathbf{r}')] d^3\mathbf{k} d^3\mathbf{k}' = \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} G_{jk}^{(j)}(\omega, \mathbf{k}) \delta(\mathbf{k} - \mathbf{k}') \exp[i(\mathbf{k}\mathbf{r} - \mathbf{k}'\mathbf{r}')] d^3\mathbf{k} d^3\mathbf{k}' = \\ &= \int_{-\infty}^{+\infty} G_{jk}^{(j)}(\omega, \mathbf{k}) \exp(i\mathbf{k}\rho) d^3\mathbf{k} \quad (\rho = \mathbf{r} - \mathbf{r}'). \end{aligned} \quad (23.13)$$



Применим равенство (23.11) к уравнениям Максвелла (23.8). Исключив при помощи материальных уравнений (23.9) индукции  $\mathbf{D}$  и  $\mathbf{B}$ , получаем

$$ik_0 \epsilon \mathbf{E} + i[\kappa \mathbf{H}] = \frac{4\pi}{c} \mathbf{j}_e, \quad ik_0 \mu \mathbf{H} - i[\kappa \mathbf{E}] = \frac{4\pi}{c} \mathbf{j}_m. \quad (23.14)$$

Для того чтобы воспользоваться формулой (23.11), надо разрешить эти уравнения относительно  $\mathbf{E}$  и  $\mathbf{H}$ . Детерминант уравнений (23.14) равен  $\mathcal{D} = k_0^2 \epsilon \mu \Delta^2$ , где

$$\Delta = \kappa^2 - k_0^2 \epsilon \mu. \quad (23.15)$$

Пусть среда однородна. Так как  $\mu(\omega, \kappa)$  не может обращаться в нуль, дисперсионное уравнение  $\mathcal{D} = 0$  распадается на два: либо  $\epsilon(\omega, \kappa) = 0$  (продольные волны), либо  $\Delta(\omega, \kappa) = 0$  (поперечные волны). Решение уравнений (23.14) есть

$$\begin{aligned} \mathbf{E} &= \frac{4\pi}{i\omega\Delta} \left\{ -k_0^2 \mu \mathbf{j}_e + \frac{\kappa}{\epsilon} (\kappa \mathbf{j}_e) + k_0 [\kappa \mathbf{j}_m] \right\}, \\ \mathbf{H} &= \frac{4\pi}{i\omega\Delta} \left\{ -k_0^2 \epsilon \mathbf{j}_m + \frac{\kappa}{\mu} (\kappa \mathbf{j}_m) - k_0 [\kappa \mathbf{j}_e] \right\}, \end{aligned}$$

или в компонентах

$$\begin{aligned} E_\alpha &= -\frac{1}{i\omega} (A_{\alpha\beta}^{(E, e)} j_{e\beta} + A_{\alpha\beta}^{(E, m)} j_{m\beta}), \\ H_\alpha &= -\frac{1}{i\omega} (A_{\alpha\beta}^{(H, e)} j_{e\beta} + A_{\alpha\beta}^{(H, m)} j_{m\beta}), \end{aligned} \quad (23.16)$$

где

$$\left[ \begin{array}{c|c} A_{\alpha\beta}^{(E, e)} & A_{\alpha\beta}^{(E, m)} \\ \hline A_{\alpha\beta}^{(H, e)} & A_{\alpha\beta}^{(H, m)} \end{array} \right] = \frac{4\pi}{\Delta} \left[ \begin{array}{c|c} k_0^2 \mu \delta_{\alpha\beta} - \frac{\kappa_\alpha \kappa_\beta}{\epsilon} & -k_0 \epsilon_{\alpha\gamma\beta} \kappa_\gamma \\ \hline k_0 \epsilon_{\alpha\gamma\beta} \kappa_\gamma & k_0^2 \epsilon \delta_{\alpha\beta} - \frac{\kappa_\alpha \kappa_\beta}{\mu} \end{array} \right]. \quad (23.17)$$

Здесь  $\epsilon_{\alpha\gamma\beta}$  — полностью антисимметричный единичный тензор третьего ранга ( $\epsilon_{123} = 1$ ), а по дважды входящему в (23.16) индексу  $\beta$ , конечно, производится суммирование. Заметим, что в силу (23.15)

$$\frac{1}{\Delta \epsilon} = \frac{1}{\kappa^2 \epsilon} + \frac{k_0^2 \mu}{\Delta \kappa^2}, \quad \frac{1}{\Delta \mu} = \frac{1}{\kappa^2 \mu} + \frac{k_0^2 \epsilon}{\Delta \kappa^2}.$$

Поэтому элементы диагональных квадратов матрицы (23.17) можно записать и в другом виде:

$$\begin{aligned} A_{\alpha\beta}^{(E, e)} &= 4\pi \left[ \frac{k_0^2 \mu}{\Delta} \left( \delta_{\alpha\beta} - \frac{\kappa_\alpha \kappa_\beta}{\kappa^2} \right) - \frac{\kappa_\alpha \kappa_\beta}{\kappa^2 \epsilon} \right], \\ A_{\alpha\beta}^{(H, m)} &= 4\pi \left[ \frac{k_0^2 \epsilon}{\Delta} \left( \delta_{\alpha\beta} - \frac{\kappa_\alpha \kappa_\beta}{\kappa^2} \right) - \frac{\kappa_\alpha \kappa_\beta}{\kappa^2 \mu} \right]. \end{aligned}$$

Согласно § 16 величины  $\mathbf{D}/4\pi = -\mathbf{j}_e/i\omega$  и  $\mathbf{B}/4\pi = -\mathbf{j}_m/i\omega$  являются для полей  $\mathbf{E}$  и  $\mathbf{H}$  обобщенными силами, так что введенные в (23.16) коэффициенты  $A$ —это именно те коэффициенты  $A_{jk}$ , которые входят в формулу (23.11). Применяя эту формулу, получаем для  $\omega\mathbf{k}$ -плотностей  $G_{\alpha\beta}^{(E)}(\omega, \mathbf{k}) = \langle E_\alpha(\omega, \mathbf{k}) E_\beta^*(\omega, \mathbf{k}) \rangle$  и т. д. следующие выражения:

$$G_{\alpha\beta}^{(E)}(\omega, \mathbf{k}) = -\frac{i\Theta(\omega, T)}{4\pi^3\omega} \left\{ \frac{k_0^2}{\Delta} \left( \delta_{\alpha\beta} - \frac{x_\alpha x_\beta}{k^2} \right) - \frac{x_\alpha x_\beta}{\varepsilon k^2} - \text{к. с.} \right\}, \quad (23.18a)$$

$$G_{\alpha\beta}^{(H)}(\omega, \mathbf{k}) = -\frac{i\Theta(\omega, T)}{4\pi^3\omega} \left\{ \frac{k_0^2}{\Delta} \left( \delta_{\alpha\beta} - \frac{x_\alpha x_\beta}{k^2} \right) - \frac{x_\alpha x_\beta}{\mu k^2} - \text{к. с.} \right\}, \quad (23.18б)$$

$$G_{\alpha\beta}^{(EH)}(\omega, \mathbf{k}) = -G_{\alpha\beta}^{(HE)}(\omega, \mathbf{k}) = \frac{i\Theta(\omega, T)}{4\pi^3 c} \varepsilon_{\alpha\gamma\beta} k_\gamma \left( \frac{1}{\Delta} - \frac{1}{\Delta^*} \right). \quad (23.18в)$$

Из (23.18в) мы видим, что, в отличие от сторонних токов  $\mathbf{j}_e$  и  $\mathbf{j}_m$ , компоненты  $\mathbf{E}$  и  $\mathbf{H}$  взаимно коррелированы, но только их *ортогональные* компоненты, так как  $\varepsilon_{\alpha\gamma\beta} = 0$  при  $\alpha = \beta$ . Первые члены в формулах (23.18а, б), содержащиеся в знаменателях  $\Delta$ , определяют корреляцию поперечного (по отношению к  $\mathbf{k}$ ) поля, т. е. напряженностей  $\mathbf{E}_\perp = \mathbf{E} - \mathbf{E}_\parallel$  и  $\mathbf{H}_\perp = \mathbf{H} - \mathbf{H}_\parallel$ , а вторые члены—корреляцию продольного поля  $\mathbf{E}_\parallel = \mathbf{x}(\mathbf{x}\mathbf{E})/k^2$ , и аналогично для  $\mathbf{H}_\parallel$ . Между собой поперечные и продольные напряженности не коррелированы (в силу чего  $G = G_\perp + G_\parallel$ ).

Остановимся в заключение настоящего параграфа на флуктуациях *полного* тока в среде, под которым понимается ток, состоящий из всех «видов» электрического тока, за исключением вакуумного тока смещения. В рамках микроскопической теории полный ток—это сумма конвекционных токов, обусловленных движением любых (не подразделяемых на «виды») микрозарядов, усредненная по физически бесконечно малым объемам среды. При таком усреднении лоренцевых уравнений для микрополей  $\mathbf{e}$  и  $\mathbf{h}$  мы получаем для  $\mathbf{E} = \bar{\mathbf{e}}$  и  $\mathbf{B} = \bar{\mathbf{h}}$  (волнистая черта—усреднение по физически бесконечно малому объему) макроскопические уравнения следующего вида:

$$\text{rot } \mathbf{B}' = \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} + \frac{4\pi}{c} \mathbf{j}_n, \quad \text{rot } \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{B}'}{\partial t}. \quad (23.19)$$

Здесь через  $\mathbf{j}_n$  обозначена объемная плотность полного тока, а штрих над  $\mathbf{B}$  поставлен для того, чтобы не смешивать эту магнитную индукцию, включающую и *стороннюю* магнитную индукцию, с индукцией  $\mathbf{B}$  в уравнениях (14.3), в которых *сторонний* магнитный ток выделен. Объемные плотности полного заряда и полного тока связаны, конечно, уравнением непрерывности

$$\frac{\partial \rho_n}{\partial t} + \text{div } \mathbf{j}_n = 0. \quad (23.20)$$

Именно эти полные плотности представляют интерес в ряде задач, в частности, в задачах о флуктуациях в плазме.

Установим прежде всего связь величины  $\mathbf{B}'$  и  $\mathbf{j}_a$  с величинами, фигурирующими в стандартной симметричной форме уравнений Максвелла. Запишем уравнения (23.19) и (23.20) для  $\omega$ -амплитуд:

$$i[\kappa \mathbf{B}'] = -ik_0 \mathbf{E} + \frac{4\pi}{c} \mathbf{j}_a, \quad i[\kappa \mathbf{E}] = ik_0 \mathbf{B}', \quad (23.21)$$

$$-\omega \rho_n + \kappa \mathbf{j}_a = 0. \quad (23.22)$$

Сопоставление вторых уравнений (23.21) и (23.8) показывает, что  $\mathbf{B}' = \mathbf{B} - \frac{4\pi}{ik_0 c} \mathbf{j}_a$ , в силу чего первое уравнение (23.21) принимает вид

$$i[\kappa \mathbf{B}] - \frac{4\pi}{k_0 c} [\kappa \mathbf{j}_a] = -ik_0 \mathbf{E} + \frac{4\pi}{c} \mathbf{j}_a.$$

Вычитая отсюда первое уравнение (23.8), получаем

$$\mathbf{j}_a = \mathbf{j}_e - \frac{ik_0 c}{4\pi} (\mathbf{D} - \mathbf{E}) - \frac{1}{k_0} [\kappa \mathbf{j}_a] + \frac{ic}{4\pi} [\kappa (\mathbf{B} - \mathbf{H})]. \quad (23.23)$$

Отсюда непосредственно видно, что полный ток складывается из стороннего электрического тока (первый член), наведенного поляризационного тока (второй член, включающий, конечно, и ток проводимости) и электрических токов, обусловленных сторонним магнитным током (третий член) и наведенным намагничиванием (последний член, включающий и магнитную проводимость).

Нетрудно записать полный ток в функции только электрической напряженности  $\mathbf{E}$ . Для этого достаточно подставить в (23.23) вытекающие из (23.8) выражения для сторонних токов  $\mathbf{j}_e$  и  $\mathbf{j}_a$ . Это дает

$$\mathbf{j}_a = \frac{i\omega}{4\pi} \left\{ \mathbf{E} + \frac{1}{k_0^2} [\kappa [\kappa \mathbf{E}]] \right\},$$

или в компонентах

$$j_{\alpha\beta} = \frac{i\omega}{4\pi} \left\{ (1 - \eta^2) \delta_{\alpha\beta} + \eta^2 \frac{\kappa_\alpha \kappa_\beta}{\kappa^2} \right\} E_\beta \quad \left( \eta = \frac{\kappa}{k_0} \right).$$

Следовательно,  $\omega$ -плотности полного тока  $\mathbf{j}_a$  и электрической напряженности  $\mathbf{E}$  связаны соотношением

$$G_{\alpha\beta}^{(j)}(\omega, \kappa) =$$

$$= \frac{\omega^2}{16\pi^2} \left[ (1 - \eta^2) \delta_{\alpha\gamma} + \eta^2 \frac{\kappa_\alpha \kappa_\gamma}{\kappa^2} \right] \left[ (1 - \eta^2) \delta_{\beta\delta} + \eta^2 \frac{\kappa_\beta \kappa_\delta}{\kappa^2} \right] G_{\gamma\delta}^{(E)}(\omega, \kappa).$$

Подставив сюда выражение (23.18a) для  $G_{\gamma\delta}^{(E)}$  и выполнив суммирование по дважды входящим в правую часть индексам  $\gamma$  и  $\delta$ ,

находим

$$G_{\alpha\beta}^{(j_p)}(\omega, \kappa) = -\frac{i\omega\Theta(\omega, T)}{64\pi^3} \left\{ \frac{k_{\alpha\beta}^2}{\Delta} (1 - \eta^2)^2 \left( \delta_{\alpha\beta} - \frac{\kappa_{\alpha}\kappa_{\beta}}{\kappa^2} \right) - \frac{\kappa_{\alpha}\kappa_{\beta}}{\varepsilon\kappa^2} - \text{к. с.} \right\}. \quad (23.24)$$

Для  $\omega\kappa$ -плотности флуктуаций полного заряда  $\rho_n$  единицы объема среды, в соответствии с уравнением непрерывности (23.22) и результатом (23.24), получаем

$$G^{(p_n)}(\omega, \kappa) = \frac{\kappa_{\alpha}\kappa_{\beta}}{\omega^2} G_{\alpha\beta}^{(j_p)}(\omega, \kappa) = \frac{i\Theta(\omega, T)\kappa^2}{64\pi^3\omega} \left( \frac{1}{\varepsilon} - \frac{1}{\varepsilon^*} \right). \quad (23.25)$$

В флуктуации полного тока (23.24) вносят вклад как поперечные, так и продольные волны, но флуктуации объемной плотности полного заряда (23.25) связаны только с продольными волнами: в знаменателе выражения для  $G^{(p_n)}$  содержится только  $\varepsilon(\omega, \kappa)$ , а  $\Delta(\omega, \kappa)$  не входит. Множитель  $\kappa^2$  в (23.25) подчеркивает большую роль мелкомасштабных флуктуаций. Из-за того, что в однородной среде  $\varepsilon$  — функция только  $\kappa = |\kappa|$ , преобразование Фурье (23.13), дающее спектральную плотность флуктуаций, принимает для заряда  $\rho_n$  вид

$$\begin{aligned} g^{(p_n)}(\omega, \rho) &= \frac{i\Theta(\omega, T)}{64\pi^3\omega} \int_{-\infty}^{+\infty} \exp(i\kappa\rho) \left( \frac{1}{\varepsilon} - \frac{1}{\varepsilon^*} \right) \kappa^2 d^3\kappa = \\ &= \frac{i\Theta(\omega, T)}{16\pi^3\omega\rho} \int_0^{\infty} \left( \frac{1}{\varepsilon} - \frac{1}{\varepsilon^*} \right) \kappa^2 \sin \kappa\rho d\kappa \quad (\rho = |\mathbf{r} - \mathbf{r}'|). \end{aligned} \quad (23.26)$$

В задаче 7 формулы (23.24) и (23.25) применены к простой модели среды, обладающей пространственной дисперсией и представляющей интерес для описания флуктуаций в ионосферной плазме.

Сделаем несколько дополнительных замечаний в заключение этой главы.

В § 20 мы отметили, что теория случайных полей, обусловленных равновесными тепловыми флуктуациями, относится к статистической схеме 1), т. е. имеет дело с задачами, в которых известна статистика случайных источников поля. Тем не менее мы выделили тепловые поля в отдельную главу, указав не только на их важность, но и на специфичность. Последняя заключается именно в том, что в данном круге задач статистика источников не задается извне, а *определяется самими динамическими уравнениями задачи*. Говоря точнее, функции корреляции сторонних (ланжевековских) сил определены антиэрмитовыми частями тех детерминированных линейных операторов, через которые записываются динамические уравнения. В этом и состоит флуктуационно-диссипационная теорема (§ 15).

Если же речь идет о существенно неравновесных тепловых флуктуациях или вообще о полях нетеплового происхождения, то статистика источников обычно не определена видом самих уравнений, а должна быть задана — либо формально, либо на основе статистического исследования микроскопических моделей.

Другое замечание касается универсальности различных форм ФДТ. Мы рассматривали в этой главе *электромагнитные* тепловые поля, хотя опирались на ФДТ, которая применима к равновесным тепловым флуктуациям *любой* физической природы (§ 15). Для электромагнитных полей были получены другие формы ФДТ — кирхгофская (§§ 17, 18), волноводная (§ 19), равновесная (§ 21), — более прозрачные физически, более «экономные» по процедуре решения конкретных задач и обладающие более широкими возможностями в отношении охвата задач, поддающихся решению. Естественно возникает вопрос о том, являются ли эти формы ФДТ столь же физически универсальными, как и исходная «каноническая» ее форма (15.14).

Все, что нам понадобилось учесть в электродинамике для получения указанных форм ФДТ, сводится к электродинамической теореме взаимности (17.2) или (22.1) и комплексной лемме Лоренца (21.2). Можно поэтому ожидать, что для любых равновесных (и квазиравновесных) флуктуационных полей, для которых справедливы теорема взаимности и аналог леммы Лоренца, будут верны и перечисленные формы ФДТ. Анализ этого вопроса показывает, что дело обстоит именно так [15], и отсюда вытекает ряд интересных следствий.

Например, кирхгофская форма ФДТ и, в частности, формула (18.15) полностью справедливы для акустических волн. Тепловые потери обусловлены здесь вязкостью и теплопроводностью среды. Представим себе, что тело, обладающее такими потерями, погружено в «прозрачную» жидкость, т. е. жидкость с малым поглощением продольных волн. Послав на тело плоскую волну частоты  $\omega$  (скажем, в звуковом или ультразвуковом диапазоне), мы можем найти эффективный поперечник поглощения тела  $\sigma_{\text{эфф}}$  (при данной ориентации тела). Формула (18.15), в которой можно, конечно, положить  $\Theta = kT$ , определяет тогда поток энергии продольных волн частоты  $\omega$ , излучаемых телом в телесный угол  $d\Omega$ , в направлении, противоположном направлению прихода вспомогательной плоской волны. Хотя физический механизм тепловых флуктуаций в теле (флуктуаций плотности, температуры, скоростей макрочастиц) совершенно иной, чем в электродинамике, «вызвучивание» нагретого тела в окружающую «прозрачную» среду происходит по такому же закону, как и «высвечивание» на электромагнитных волнах, а именно по закону Кирхгофа.

## Задачи

1. С учетом реакции излучения уравнение движения упруго связанного электрона имеет при малых его скоростях  $|\dot{r}| \ll c$  вид

$$m\ddot{r} - \frac{2e^2}{3c^3} \ddot{r} + m\omega_0^2 r = I(t) = eE(t), \quad (1)$$

где  $e$  и  $m$  — заряд и масса электрона,  $c$  — скорость света в вакууме,  $\omega_0$  — собственная частота осциллятора. Пусть  $E(t)$  — напряженность электрического поля равновесного излучения. Применяя ФДТ к уравнению (1), найти спектральную плотность энергии этого излучения (по положительным частотам  $\omega$ ), т. е. величину (электрон находится в вакууме)

$$u_{0\omega} = \frac{1}{8\pi} \langle |E|^2 + |H|^2 \rangle.$$

Решение. Уравнение движения (1) в спектральной форме есть

$$r(\omega) = \alpha(\omega) I(\omega),$$

где

$$\frac{1}{\alpha(\omega)} = m(\omega_0^2 - \omega^2) - i \frac{2e^2 \omega^3}{3c^3}. \quad (2)$$

Спектральная плотность компоненты  $E_x(t) = f_x(t)/e$  флуктуационного поля по  $\omega > 0$  равна

$$g_{E_x}^+(\omega) = \frac{1}{e^2} g_{f_x}^+(\omega) = \frac{2}{e^2} \langle |f_x(\omega)|^2 \rangle, \quad (3)$$

где  $\langle |f_x(\omega)|^2 \rangle$  — спектральная плотность силы  $f_x(t)$  по  $\omega$  от  $-\infty$  до  $+\infty$ . Согласно дискретной ФДТ  $\langle |f_x(\omega)|^2 \rangle$  выражается через обобщенную восприимчивость  $\alpha(\omega)$  следующим образом:

$$\langle |f_x(\omega)|^2 \rangle = - \frac{i\Theta(\omega, T)}{2\pi\omega} \left[ \frac{1}{\alpha^*(\omega)} - \frac{1}{\alpha(\omega)} \right] \quad (4)$$

(это формула (15.96) в частном случае единственной переменной). Из (2) — (4) следует, что

$$g_{E_x}^+(\omega) = \frac{4\omega^2}{3\pi c^3} \Theta(\omega, T).$$

Заметим, что  $\text{Re}\{1/\alpha(\omega)\}$  не играет роли, и поэтому электрон мог бы быть и свободным ( $\omega_0 = 0$ ).

Для того чтобы получить спектральную плотность энергии электромагнитного поля, достаточно учесть, что равновесное излучение в неограниченном пространстве изотропно, в силу чего

$$\langle |E_x|^2 \rangle = \langle |E_y|^2 \rangle = \langle |E_z|^2 \rangle,$$

и что его магнитная энергия равна электрической:

$$\langle |E|^2 \rangle = \langle |H|^2 \rangle.$$

Поэтому

$$u_{0\omega} = 6 \frac{1}{8\pi} g_{E_x}^+(\omega) = \frac{\omega^2}{\pi^2 c^3} \Theta(\omega, T),$$

т. е. мы получили формулу Планка.

2. Формулы (16.4) для корреляционных функций сторонних токов можно получить и более физическим путем, чем использованный в § 16, опираясь на классический закон Кирхгофа и принцип детального равновесия. Последний требует, чтобы для двух одинаково нагретых тел 1 и 2 мощность  $Q_{12}$ , поглощаемая телом 1 на излучения тела 2, была на каждой частоте  $\omega$  равна мощности  $Q_{21}$ , поглощаемой телом 2 из излучения тела 1. Пользуясь этим, вывести формулы (16.4) из рассмотрения обмена энергией между телом и пластиной (рис. 20), разнесенными на столь большое расстояние  $R$ , что их размеры гораздо меньше радиуса  $\sqrt{\lambda R}$  первой зоны Френеля (т. е. они находятся во френуговской зоне друг друга). Вместе с тем размеры пластины настолько больше длины волны  $\lambda$ , что излучение пластины можно вычислять по классическому закону Кирхгофа.

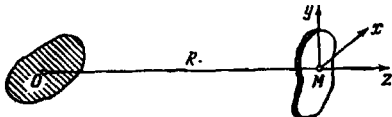


Рис. 20.

Решение. Подсчитаем при указанных условиях мощности  $Q_T$  и  $Q_0$ , поглощаемые телом и пластиной на частоте  $\omega$ .

Пусть  $A$  — коэффициент поглощения пластины и  $\Sigma$  — ее площадь ( $\lambda R \gg \Sigma \gg \lambda^2$ ). Поглощаемая пластиной мощность есть

$$Q_0 = \mathcal{S}_z A \Sigma, \quad (1)$$

где  $\mathcal{S}_z$  —  $z$ -компонента вектора Пойнтинга теплового излучения тела. Для  $\mathcal{S}_z$  на одной поляризации мы имеем выражение (18.13), так что при учете обеих независимых поляризаций  $\mathcal{S}_z = \frac{c}{\pi} \langle |E_1|^2 \rangle$ . Подставив сюда  $E_1$  из (18.12) и записывая скалярные произведения векторов в компонентах, получаем в соответствии с (1), что

$$Q_0 = \frac{c A \Sigma}{\pi} \iint \{ E_{0\alpha}(r) E_{0\beta}^*(r') \langle j_{e\alpha}(r) j_{e\beta}^*(r') \rangle + \\ + H_{0\alpha}(r) H_{0\beta}^*(r') \langle j_{m\alpha}(r) j_{m\beta}^*(r') \rangle - E_{0\alpha}(r) H_{0\beta}^*(r') \langle j_{e\alpha}(r) j_{m\beta}^*(r') \rangle - \\ - H_{0\alpha}(r) E_{0\beta}^*(r') \langle j_{m\alpha}(r) j_{e\beta}^*(r') \rangle \} d^3r d^3r'. \quad (2)$$

Теперь подсчитаем  $Q_T$  — мощность, поглощаемую телом из излучения пластины. Так как ее размеры велики по сравнению с  $\lambda$ , для интенсивности излучения пластины в телесный угол  $d\Omega$ , ось которого направлена на тело, справедлива формула (18.16).

Сопоставим выражение (18.16) с мощностью, излучаемой в тот же телесный угол  $d\Omega$  единичным точечным источником (электрическим диполем в точке  $M$  с моментом  $p = 1/i\omega$ , лежащим в плоскости пластины):

$$dP_{0\omega} = \frac{\omega^2}{8\pi c^3} d\Omega.$$

Мы видим, что пластина излучает в

$$\frac{dP_{0\omega}}{dP_{0\omega}} = \frac{2c\Theta}{\pi^2} A \Sigma$$

раз больше точечного источника. Следовательно, во столько же раз больше и мощность  $Q_T$  по сравнению с мощностью  $Q_0$ , поглощаемой телом из

излучения точечного источника:

$$Q_1 = \frac{2c\Theta}{\pi^3} A \Sigma Q_0.$$

Потери же  $Q_0$  дифракционного поля, создаваемого электрическим диполем, равны

$$Q_0 = -\frac{i\omega}{16\pi} \int \left\{ E_{0\alpha}(r) E_{0\beta}^*(r) (\varepsilon_{\alpha\beta} - \varepsilon_{\beta\alpha}^*) + H_{0\alpha}(r) H_{0\beta}^*(r) (\mu_{\alpha\beta} - \mu_{\beta\alpha}^*) \right\} d^3r,$$

так что

$$Q_1 = -\frac{i\omega c A \Sigma \Theta(\omega, T)}{8\pi^3} \int \left\{ E_{0\alpha}(r) E_{0\beta}^*(r) (\varepsilon_{\alpha\beta} - \varepsilon_{\beta\alpha}^*) + \right. \\ \left. + H_{0\alpha}(r) H_{0\beta}^*(r) (\mu_{\alpha\beta} - \mu_{\beta\alpha}^*) \right\} d^3r. \quad (3)$$

Выражения (2) и (3) должны быть равны друг другу. Почленное сравнение подынтегральных выражений дает формулы (16.4). Однозначность результата следует из того, что равенство интегралов должно иметь место при любой форме, величине и расположении тела и пластины (конечно, не нарушающих условий  $\lambda R \gg \Sigma \gg \lambda^2$ ).

Следует заметить, что при данном выводе формул (16.4) надо понимать под  $\Theta(\omega, T)$  только ту часть средней энергии осциллятора, которая зависит от температуры. Энергия нулевых колебаний  $\hbar\omega/2$  не должна учитываться, поскольку рассматривается обмен энергией посредством излучения (§ 17).

3. В волноводе между двумя черными излучателями с температурой  $T_0$  находится тело с температурой  $T \neq T_0$ . Найти поток энергии флуктуационного поля на  $n$ -й собственной волне волновода и показать, что при любых  $T$  и  $T_0$  нулевые колебания из этого потока выпадают.

Решение. Поток энергии, скажем, справа от тела равен собственному излучению тела ( $P_{\omega n}^+$ ) плюс отраженное телом излучение правого излучателя  $\left( \frac{\Theta_0}{2\pi} \sum_m' R_{mn}^- \right)$  плюс прошедшее излучение левого излучателя  $\left( \frac{\Theta_0}{2\pi} \sum_m' D_{mn}^+ \right)$  и минус встречный поток излучения справа ( $\Theta_0/2\pi$ ):

$$P_{\omega n} = P_{\omega n}^+ + \frac{\Theta_0}{2\pi} \sum_m' (R_{mn}^- + D_{mn}^+) - \frac{\Theta_0}{2\pi}.$$

Согласно (19.3)

$$P_{\omega n}^+ = \frac{\Theta}{2\pi} \left[ 1 - \sum_m' (R_{mn}^- + D_{mn}^+) \right],$$

так что

$$P_{\omega n} = \frac{\Theta - \Theta_0}{2\pi} \left[ \sum_m' (R_{mn}^- + D_{mn}^+) - 1 \right].$$

Но нулевая энергия  $\hbar\omega/2$  из разности  $\Theta - \Theta_0$  выпадает, т. е. встречные потоки энергии нулевых колебаний взаимно уничтожаются при любых  $T$  и  $T_0$ . Именно потому, что эти колебания никогда не участвуют в переносе энергии, член  $\hbar\omega/2$  в  $\Theta(\omega, T)$  можно отбрасывать во всех случаях, когда речь идет о (среднем) потоке энергии.



4. Найти мощность, излучаемую в идеальный волновод на волне с постоянной распространения  $h$  излучателем, состоящим из поперечной идеально отражающей перегородки и поперечной же полупрозрачной пластинки, отстоящей от перегородки на расстояние  $l$  (рис. 21). Пластика предполагается тонкой, т. е. ее толщина мала по сравнению с длиной волны в ее материале, так что интерференции внутри пластинки можно не учитывать и характеризовать пластинку вещественными амплитудными коэффициентами отражения ( $r$ ) и пропускания ( $d$ ). При этом  $r^2 + d^2 + a = 1$ , где  $a$  — энергетический коэффициент поглощения пластинки.

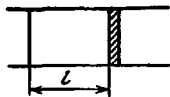


Рис. 21.

Решение. Излучаемая мощность  $P = \frac{\Theta}{2\pi} A$ , где

$A = 1 - \mathcal{R}$  (излучатель в целом непрозрачен),  $\mathcal{R}$  — энергетический коэффициент отражения излучателя. Если справа падает волна  $e^{-ihz}$ , то в результате первого отражения от пластинки и последующих многократных повторных отражений прошедшей через нее волны между пластинкой и идеально отражающей перегородкой, амплитуда отраженной излучателем волны будет

$$r + d^2 e^{-2ihl} + r d^2 e^{-4ihl} + r^2 d^2 e^{-6ihl} + \dots = r + \frac{d^2}{e^{2ihl} - r}.$$

Следовательно,

$$A = 1 - \mathcal{R} = 1 - \left| r + \frac{d^2}{e^{2ihl} - r} \right|^2 = a \left[ 1 + \frac{d^2}{1 + r^2 - 2r \cos 2hl} \right].$$

От  $l$  поглощение зависит осцилляторно с периодом, равным половине длины волны  $\Lambda = 2\pi/h$ , и с максимумами на резонансных длинах резонатора, образуемого перегородкой и пластинкой. При уменьшении прозрачности пластинки ( $d \rightarrow 0$ ) получаем  $A = a$ . Если же коэффициент отражения от пластинки  $r^2 \rightarrow 1$  (при этом с необходимостью  $d \rightarrow 0$ ,  $a \rightarrow 0$ ), то  $A \rightarrow 0$ .

5. Формула Найквиста для спектральной плотности тепловой э. д. с.  $e(t)$  в двуполоснике с импедансом  $Z(\omega)$  имеет вид

$$\langle |\tilde{e}(\omega)|^2 \rangle = \frac{1}{\pi} \Theta(\omega, T) \operatorname{Re} Z(\omega).$$

Для хороших проводников часто заменяют  $\operatorname{Re} Z$  на омическое сопротивление  $R$ . При каких условиях эта замена законна? Ведь при переходе от проводника к идеальному диэлектрику (проводимость  $\sigma \rightarrow 0$  и соответственно  $R \rightarrow \infty$ ) получается неограниченное нарастание э. д. с. Как уже было отмечено в ч. 1, § 54, этот кажущийся парадокс возникает при забвении того, что замена  $\operatorname{Re} Z$  на  $R$  законна лишь для столь хороших проводников, внутри которых можно пренебречь током смещения по сравнению с током проводимости. Ясно, что при  $\sigma \rightarrow 0$  это условие рано или поздно нарушается. Вывести формулу для  $\langle |\tilde{e}(\omega)|^2 \rangle$  в тонком проводе с учетом тока смещения.

Решение. Если при выводе формулы (20.9) для  $\langle |I(s)|^2 \rangle$ , а значит, и при выводе выражения (20.13) для  $\langle |\tilde{e}(s)|^2 \rangle$  не пренебрегать током смещения в проводе, т. е. пользоваться общим выражением (20.5) для корреляции компонент теплового поля, то результатом снова будут формулы (20.9) и (20.13), но, в отличие от (20.10), теперь

$$dQ_0 = \frac{|j_0|^2}{2\sigma} \frac{d^2 r}{1 + (\varepsilon\omega/4\pi\sigma)^2}.$$

Следовательно, для квазистационарной ( $j_0$  не зависит от  $s$ ) равновесной ( $\Theta = \text{const}$ ) цепи энергетическое сопротивление, определенное равенством,

аналогичным (20.15), равно

$$R_s = \frac{1}{|I_0|^2} \int \frac{|j_0|^2 d^3r}{\sigma [1 + (e'\omega/4\pi\sigma)^2]}. \quad (1)$$

Смысл этого выражения лучше всего пояснить переходом к предельному случаю слабого скин-эффекта в проводе из однородного материала ( $e'$  и  $\sigma$  постоянны). Тогда  $j_0 = I_0/\Sigma$  и для участка провода длины  $l$  формула (1) принимает вид

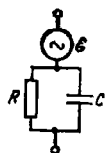


Рис. 22.

$$R_s = \frac{R}{1 + (RC\omega)^2},$$

где  $R = l/\sigma\Sigma$  — омическое сопротивление, а  $C = e'\Sigma/4\pi l$  — емкостная характеристика данного участка провода. Таким образом,  $R_s$  представляет собой в этом случае активное сопротивление  $RC$ -ячейки (рис. 22) и при  $R \rightarrow \infty$  получаем  $R_s = 0$ . Замена  $R_s$  на  $R$  справедлива лишь при условии  $(RC\omega)^2 \ll 1$ .

6. Пользуясь формулами (15.11а, б), получить корреляционные функции  $\omega$ -амплитуд полей  $\xi(t, r)$  и  $f(t, r)$  в неограниченной однородной среде, т. е. амплитуд в пространственных разложениях Фурье

$$\xi(\omega, r) = \int_{-\infty}^{+\infty} \xi(\omega, k) \exp(i\mathbf{k}r) d^3k, \quad f(\omega, r) = \int_{-\infty}^{+\infty} f(\omega, k) \exp(i\mathbf{k}r) d^3k.$$

Решение. Проведем сначала расчет функции корреляции для  $\omega$ -амплитуд  $f(t, r)$ :

$$f(\omega, k) = \frac{1}{(2\pi)^3} \int_{-\infty}^{+\infty} f(\omega, r) \exp(-i\mathbf{k}r) d^3r,$$

С учетом (15.11б) имеем

$$\begin{aligned} \langle f(\omega, k) f^*(\omega, k') \rangle &= \\ &= \frac{1}{(2\pi)^6} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \langle f(\omega, r) f^*(\omega, r') \rangle \exp(-i\mathbf{k}r + i\mathbf{k}'r') d^3r d^3r' = \\ &= -\frac{i\Theta(\omega, T)}{(2\pi)^3 \omega} \int_{-\infty}^{+\infty} (\hat{A}^{-1} - \hat{A}) \delta(r-r') \exp(-i\mathbf{k}r + i\mathbf{k}'r') d^3r d^3r'. \end{aligned} \quad (1)$$

Представим дельта-функцию в виде интеграла Фурье:

$$\delta(r-r') = \frac{1}{(2\pi)^3} \int_{-\infty}^{+\infty} e^{ip(r-r')} d^3p.$$

Поскольку  $\nabla e^{ipr} = ip e^{ipr}$ , действие линейного оператора  $\hat{A}^{-1}(\nabla)$  на  $e^{ipr}$  дает

$$\hat{A}^{-1}(\nabla) \delta(r-r') = \frac{1}{(2\pi)^3} \int_{-\infty}^{+\infty} A^{-1}(ip) e^{ip(r-r')} d^3p,$$

где  $A^{-1}(ip)$  — функция  $p$ , получающаяся при замене  $\nabla \rightarrow ip$  в операторе  $\hat{A}^{-1}(\nabla)$ . Аналогично <sup>1)</sup>,

$$\hat{A}^{-1*}(\nabla') \delta(r-r') = \frac{1}{(2\pi)^3} \int_{-\infty}^{+\infty} A^{-1*}(ip) e^{ip(r-r')} d^3p,$$

так что (1) принимает вид

$$\begin{aligned} \langle f(\omega, \kappa) f^*(\omega, \kappa') \rangle = & -\frac{i\Theta(\omega, T)}{(2\pi)^{10} \omega} \int_{-\infty}^{+\infty} \{A^{-1*}(ip) - A^{-1}(ip)\} d^3p \times \\ & \times \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \exp\{i[(p-\kappa)r - (p-\kappa')r']\} d^3r d^3r'. \end{aligned}$$

Интегралы по  $r$  и  $r'$  дают  $(2\pi)^4 \delta(p-\kappa) \delta(p-\kappa')$ , после чего интегрирование по  $p$  приводит к искомому результату:

$$\langle f(\omega, \kappa) f^*(\omega, \kappa') \rangle = -\frac{i\Theta(\omega, T)}{(2\pi)^4 \omega} \{A^{-1*}(i\kappa) - A^{-1}(i\kappa)\} \delta(\kappa - \kappa'). \quad (2)$$

Можно получить функцию корреляции для  $\xi(\omega, \kappa)$  таким же путем (пользуясь формулой (15.1а)), а можно сделать это и иначе — исходя из того, что  $\xi(\omega, \kappa) = A(-i\omega, i\kappa) f(\omega, \kappa)$ . Полагая  $\kappa' = \kappa$  в множителе при  $\delta(\kappa - \kappa')$ , имеем

$$\begin{aligned} \langle \xi(\omega, \kappa) \xi^*(\omega, \kappa') \rangle &= A(i\kappa) A^*(i\kappa') \langle f(\omega, \kappa) f^*(\omega, \kappa') \rangle = \\ &= -\frac{i\Theta(\omega, T)}{(2\pi)^4 \omega} A(i\kappa) A^*(i\kappa) \{A^{-1*}(i\kappa) - A^{-1}(i\kappa)\} \delta(\kappa - \kappa') = \\ &= -\frac{i\Theta(\omega, T)}{(2\pi)^4 \omega} \{A(i\kappa) - A^*(i\kappa)\} \delta(\kappa - \kappa'). \quad (3) \end{aligned}$$

Уравнения (15.1) (если явным образом записать зависимость операторов  $\hat{A}$  и  $\hat{A}^{-1}$  от  $-i\omega$ ) принимают для  $\omega$ -амплитуд вид

$$\xi(\omega, \kappa) = A(-i\omega, i\kappa) f(\omega, \kappa), \quad f(\omega, \kappa) = A^{-1}(-i\omega, i\kappa) \xi(\omega, \kappa).$$

Первым из них мы только что воспользовались при выводе (3), а из второго видно, что уравнение

$$A^{-1}(-i\omega, i\kappa) = 0 \quad (4)$$

<sup>1)</sup> В связи с этой и последующими формулами необходимо сделать одно разъяснение. Знак комплексного сопряжения  $*$  понимается здесь и далее в обычном смысле, т. е. как замена  $i$  на  $-i$  во всех аргументах вещественной функции от  $-i\omega$ ,  $i\kappa$ ,  $ip$  и т. д. В книге [6] использовано другое (и менее удачное) обозначение: знак  $*$  означает там комплексное сопряжение только в спектральной амплитуде, т. е. в вещественной функции от  $i\omega$  и  $\kappa$ , и сводится, таким образом, к замене  $-i\omega \rightarrow +i\omega$ . Поэтому в [6] выражение для  $\hat{A}^{-1*}(\nabla') \delta(r-r')$  содержит под интегралом функцию  $A^{-1*}(-ip)$ . Звездочка означает изменение знака в  $i\omega$ , а в аргументе  $-ip$  знак минус записан явно. Обычное обозначение, принятое в данной книге, конечно, проще.

представляет собой дисперсионное уравнение задачи. Из (3) ясно, что любой интеграл по  $\mathbf{x}$  и  $\mathbf{x}'$ , содержащий  $\langle \xi(\omega, \mathbf{x}) \xi^*(\omega, \mathbf{x}') \rangle$ , сразу же сводится к интегралу только по  $\mathbf{x}$  (из-за наличия  $\delta(\mathbf{x} - \mathbf{x}')$ ), а последний часто может быть взят вычетами в полюсах функции  $A(-i\omega, i\mathbf{x})$ , т. е. в нулях функции  $A^{-1}(i\omega, i\mathbf{x})$ , или, иначе говоря, при значениях  $\mathbf{x}$ , являющихся корнями дисперсионного уравнения (4).

Нетрудно воспроизвести расчеты, сделанные при выводе формул (2) и (3), для случая *многомерных* однородных полей  $\xi^{(j)}(t, \mathbf{r})$  и  $j^{(j)}(t, \mathbf{r})$ , используя при этом формулы (15.14а, б). Это приводит к корреляционным матрицам (23.11) и (23.12) для  $\omega\mathbf{x}$ -амплитуд.

7. Для продольных волн, длинных по сравнению с радиусом нелокальности  $a$  ( $ka \ll 1$ ), не слишком разреженная плазма, состоящая из электронов, ионов разного сорта и нейтральных атомов и молекул, может быть хорошо описана в так называемом *квазигидродинамическом* приближении. Уравнения гидродинамики надо писать при этом для всех видов частиц, но мы ограничимся простейшей моделью *одножидкостной* (электронной) плазмы, т. е. будем считать ионы неподвижными<sup>1)</sup>.

Тогда линеаризованное уравнение движения электронной жидкости (газа) запишется в виде

$$m \frac{\partial \mathbf{v}}{\partial t} + m \mathbf{v} \mathbf{v} = - \frac{1}{N} \nabla \bar{p} - e \mathbf{E}, \quad (1)$$

где  $m$  — масса электрона,  $e$  — абсолютная величина его заряда,  $N$  — средняя концентрация электронов,  $\bar{p}$  — переменная часть давления:

$$\bar{p} = \gamma \Theta \bar{N} \quad (2)$$

( $\gamma$  — отношение Пуассона,  $\Theta = kT$  — энергетическая температура электронного газа,  $\bar{N}$  — переменная часть концентрации). В уравнение (1) введено также «трение» — через эффективное число соударений в единицу времени ( $\nu$ ). Плотность электронного тока (в данной модели это и *полный* ток) в линейном приближении равна  $\mathbf{j} = -Ne\mathbf{v}$ , а переменная часть плотности заряда есть  $\rho_f = -\bar{N}e$  (средние плотности электронного и ионного зарядов скомпенсированы, т. е. плазма квазинейтральна). Поэтому линеаризованное уравнение непрерывности будет

$$\frac{\partial \bar{N}}{\partial t} + N \operatorname{div} \mathbf{v} = 0. \quad (3)$$

Найти для описанной модели плазмы спектральную плотность флуктуаций электронной концентрации.

Решение. Согласно (23.25) для нахождения  $\omega\mathbf{x}$ -плотности флуктуаций  $\rho_f$  достаточно знать продольную диэлектрическую проницаемость  $\epsilon(\omega, \mathbf{k})$ . Для  $\omega\mathbf{x}$ -амплитуд ( $\bar{N}, \bar{p}, \mathbf{v}, \mathbf{E} \sim \exp[-i(\omega t - \mathbf{k}\mathbf{r})]$ ), исключая из (1) — (3) все величины, кроме  $\mathbf{v} = -\mathbf{j}/Ne$ , получаем следующее уравнение для полного тока  $\mathbf{j}$  (в компонентах):

$$\left(1 + i \frac{\nu}{\omega}\right) \mathbf{j} \alpha - \frac{\gamma \Theta}{m \omega^2} \kappa_\alpha \kappa_\beta \mathbf{j}_\beta = \frac{i \omega_e^2}{4 \pi \omega} \mathbf{E}_\alpha,$$

где  $\omega_e$  — плазменная (электронная) частота:

$$\omega_e^2 = 4 \pi N e^2 / m.$$

<sup>1)</sup> О более общих случаях и, в частности, о *двухжидкостной* модели (электроны и один сорт ионов), причем модели *неравновесной* (температуры электронов и ионов различны), см. в [14, 16].

Удобно ввести обозначения

$$X = \omega_s^2 / \omega^2, \quad Z = v / \omega_e, \quad a = \sqrt{\gamma} d,$$

где

$$d = \frac{1}{\omega_e} \sqrt{\frac{\Theta}{m}}$$

— так называемый *дебаевский радиус* (или радиус экранирования электрического поля, см. [16]). В этих обозначениях уравнение для  $J$  принимает вид

$$(1 + iZ \sqrt{X}) J_\alpha - a^2 X \kappa_\alpha \kappa_\beta J_\beta = \frac{i\omega}{4\pi} X E_\alpha,$$

откуда

$$J_\alpha = \frac{i\omega X}{4\pi (1 + iZ \sqrt{X})} \left( E_\alpha + \frac{a^2 X \kappa_\alpha \kappa_\beta E_\beta}{1 + iZ \sqrt{X} - \kappa^2 a^2 X} \right).$$

Электрическая индукция  $D_\alpha = \epsilon_{\alpha\beta} E_\beta$  связана с полным током  $J$  соотношением  $D = E - 4\pi J / i\omega$ <sup>1)</sup>, так что

$$D_\alpha = E_\alpha - \frac{4\pi}{i\omega} J_\alpha = \left\{ \left( 1 - \frac{X}{1 + iZ \sqrt{X}} \right) \delta_{\alpha\beta} - \frac{a^2 X^2 \kappa_\alpha \kappa_\beta}{(1 + iZ \sqrt{X}) [1 + iZ \sqrt{X} - \kappa^2 a^2 X]} \right\} E_\beta.$$

Выражение в фигурных скобках — это диэлектрическая проницаемость  $\epsilon_{\alpha\beta}(\omega, \kappa)$ . Она легко преобразуется к виду

$$\epsilon_{\alpha\beta}(\omega, \kappa) = \left( 1 - \frac{X}{1 + iZ \sqrt{X}} \right) \left( \delta_{\alpha\beta} - \frac{\kappa_\alpha \kappa_\beta}{\kappa^2} \right) + \left( 1 - \frac{X}{1 + iZ \sqrt{X} - \kappa^2 a^2 X} \right) \frac{\kappa_\alpha \kappa_\beta}{\kappa^2}.$$

откуда следуют выражения для поперечной и продольной проницаемостей:

$$\epsilon_t(\omega, \kappa) = 1 - \frac{X}{1 + iZ \sqrt{X}}, \quad (4)$$

$$\epsilon_l(\omega, \kappa) \equiv \epsilon(\omega, \kappa) = 1 - \frac{X}{1 + iZ \sqrt{X} - \kappa^2 a^2 X}.$$

Таким образом, в рассматриваемой модели  $\epsilon_t$  не обладает пространственной дисперсией (не зависит от  $\kappa$ ). Пространственная же дисперсия  $\epsilon_l = \epsilon$  возникает из-за члена с  $\sqrt{\rho}$  в (1), т. е. обусловлена *упругостью* электронного газа.

Согласно выражению (4) для  $\epsilon$  имеем

$$\frac{1}{\epsilon} - \frac{1}{\epsilon^*} = - \frac{2iZX^{3/2}}{(1 - X - \kappa^2 a^2 X)^2 + Z^2 X}.$$

<sup>1)</sup> Это ясно из первого уравнения (23.21), если записать его в виде  $i[\kappa \vec{B}] = -ik_0 D$ .

Подставив это в формулы (23.25) и (23.26), получаем, что  $\omega$ -плотность флуктуаций электронной концентрации  $\tilde{N} = -\rho_{\Phi}/e$  есть

$$G(\tilde{N})(\omega, \kappa) = \frac{1}{e^2} G(\rho_{\Phi})(\omega, \kappa) = \frac{\Theta Z X^3}{32\pi^2 \omega_e e^2} \frac{\kappa^2}{(1 - X - \kappa^2 a^2 X)^2 + Z^2 X}, \quad (5)$$

а ее спектральная плотность равна

$$g^{(\tilde{N})}(\omega, \rho) = \frac{\Theta Z X^3}{8\pi^4 \omega_e e^2 \rho} \int_0^{\infty} \frac{\kappa^2 \sin \kappa \rho}{(1 - X - \kappa^2 a^2 X)^2 + Z^2 X} d\kappa. \quad (6)$$

Интеграл в (6) легко вычисляется. Из-за четности интегранда по  $\kappa$  можно взять половину значения интеграла в пределах  $\pm \infty$ . Записав далее  $\sin \kappa \rho$  в виде  $(e^{i\kappa\rho} - e^{-i\kappa\rho})/2i$ , изменим знак  $\kappa$  в интеграле с  $e^{-i\kappa\rho}$ . Это просто удваивает первый интеграл (с  $e^{i\kappa\rho}$ ) и дает

$$g^{(\tilde{N})}(\omega, \rho) = \frac{\Theta Z X^3}{16\pi^4 i \omega_e e^2 \rho} \int_{-\infty}^{+\infty} \frac{e^{i\kappa\rho} \kappa^3 d\kappa}{(1 - X - \kappa^2 a^2 X)^2 + Z^2 X}.$$

Замыкание пути интегрирования в верхней полуплоскости комплексного  $\kappa$  сводит интеграл к вычетам в полюсах  $\kappa = k$  и  $\kappa = -k^*$ , где  $k$  — корень дисперсионного уравнения  $v(\omega, \kappa) = 0$ , лежащий в первом квадранте плоскости  $\kappa$ :  $k = k' + ik''$ .

$$= \frac{1}{a\sqrt{2X}} \left\{ \sqrt{V(1-X)^2 + Z^2 X} + (1-X) + i \sqrt{V(1-X)^2 + Z^2 X} - (1-X) \right\}. \quad (7)$$

Окончательный результат:

$$g^{(\tilde{N})}(\omega, \rho) = \frac{Nmv}{4\pi^2 \gamma^2 \Theta} \frac{e^{-k''\rho}}{\rho} \left[ \cos k'\rho + \frac{1-X}{Z\sqrt{X}} \sin k'\rho \right]. \quad (8)$$

Как это видно уже из (6),  $g^{(\tilde{N})}(\omega, 0) = \infty$ . Эта особенность при  $\rho = 0$ , вполне очевидная в (8), обусловлена недостаточно быстрым убыванием  $1/e - 1/e^*$  к нулю при возрастании  $\kappa$ . Однако само гидродинамическое описание справедливо лишь для  $\kappa a \ll 1$ , т. е. для пространственных гармоник флуктуаций  $\tilde{N}$  с длинами волн  $2\pi/\kappa \equiv \Lambda \gg a$ . Поэтому брать значений  $\rho \lesssim a$  нет смысла. В этой области  $\kappa$  фазовая скорость продольных волн становится одного порядка с тепловой скоростью электронов и, как показывает кинетическое рассмотрение, происходит сильное затухание продольных волн. Гидродинамическая модель не учитывает этого затухания, обусловленного тепловым движением зарядов ([14], § 2). Отметим в связи со сказанным, что, например, в слое  $F$  ионосферы, если принять  $N = 5 \cdot 10^8 \text{ см}^{-3}$ ,  $T = 300 \text{ К}$  и  $\gamma = 3^1$ , то

$$\omega_e = 4 \cdot 10^7 \text{ Гц}, \quad a \approx 0,3 \text{ см}.$$

В диапазоне частот  $\omega > \omega_e$  ( $X < 1$ ) и при  $Z \ll 1 - X$  из (7) имеем приближенно

$$k' \approx \frac{1}{a} \sqrt{\frac{1-X}{X}}, \quad k'' \approx \frac{Z}{2a\sqrt{1-X}}.$$

<sup>1</sup>) Основанием для такого выбора  $\gamma$  является кинетическое рассмотрение ([14], § 2).

т. е.  $k' \gg k''$ , и основную роль в (8) играет член с  $\sin k'r$ . Спектральная плотность (8) представляет собой в функции от  $r$  медленно затухающее колебание. Напротив, при  $\omega < \omega_e$  ( $X > 1$ ) и  $Z \ll X-1$  значения  $k'$  и  $k''$  меняются местами:

$$k' \approx \frac{Z}{2a \sqrt{X-1}}, \quad k'' \approx \frac{1}{a} \sqrt{\frac{X-1}{X}}$$

так что экспонента затухает гораздо быстрее, чем за период колебания  $2\pi/k'$ . В окрестности  $\omega_e$ , т. е. при  $X \approx 1$  и  $|1-X| \ll Z$ , имеем

$$k' \approx k'' \approx \frac{1}{a} \sqrt{\frac{Z}{2}}.$$

Заметим в заключение, что, ограничившись в рамках гидродинамической модели описанием состояния электронного газа переменными  $v$ ,  $\bar{\rho}$  и  $\bar{N}$ , мы исключили тем самым из рассмотрения флуктуации его температуры  $\bar{T} = T - \bar{T}$ .

## Г л а в а IV

### ТЕОРИЯ ОДНОКРАТНОГО РАССЕЯНИЯ ВОЛН

#### § 24. Метод малых возмущений

В этой и последующих главах мы займемся флуктуациями волн, распространяющихся в случайно-неоднородных средах, т. е. по классификации § 8 задачами типа 2).

Случайные неоднородности реальных сред влияют на характеристики волн, распространяющихся в этих средах, и возникающие при этом явления чрезвычайно разнообразны. Мерцание звезд и флуктуации радиоизлучения от внеземных источников, замирания (федниги) радиоволн и релеевское рассеяние света, ушнрение лазерных пучков в тропосфере и рассеяние звука в море — это лишь немногие примеры наблюдаемых эффектов. Исследованием такого рода эффектов занимается статистическая теория распространения и рассеяния волн.

Задачи о распространении волн в средах с флуктуирующими параметрами решаются, как правило, *приближенными* методами. Дело в том, что соответствующие дифференциальные уравнения содержат в коэффициентах случайные функции точки (а возможно, и времени), описывающие неоднородную среду. *Точное* решение такой *параметрической* задачи означало бы, что мы в состоянии написать, например, функцию Грина для любых реализаций входящих в уравнения случайных функций, что практически никогда не осуществимо<sup>1)</sup>. Это и вынуждает обращаться к приближенным методам. Характер приближения зависит, разумеется, от постановки задачи — слабо или сильно флуктуируют параметры среды, каково соотношение между длиной волны и размерами неоднородностей, какова геометрия задачи (длина трассы, ширина волнового пучка) и т. д. При всем разнообразии конкретных

<sup>1)</sup> Если не говорить о тех весьма частных случаях, когда среда описывается специального вида детерминированными функциями, зависящими от конечного числа случайных параметров.



условий значительная часть задач типа 2) может быть решена при помощи небольшого числа разработанных к настоящему времени приближенных методов.

Если относительные флуктуации параметров среды достаточно слабы, а рассеянное поле мало по сравнению с полем первичной волны, то применяется *метод малых возмущений*. Анализ полей, рассчитанных в *первом* порядке теории возмущений, составляет содержание *теории однократного рассеяния*, которой и посвящена данная глава.

При нарушении условий применимости теории однократного рассеяния (флуктуации в среде недостаточно слабы, рассеянное поле не мало) необходимо принимать во внимание двух-, трех- и т. д. кратное рассеяние поля, т. е. нужно строить теорию с учетом *многократного* рассеяния волн. В случае слабых, но крупных (по сравнению с длиной волны) неоднородностей многократно рассеянные волны лишь незначительно уклоняются от направления распространения первичной волны. В таких условиях многократное рассеяние эффективно описывается *методом геометрической оптики* (МГО) и примыкающими к нему более общими коротко-волновыми асимптотическими методами теории дифракции — *методом плавных возмущений* (МПВ) и *методом параболического уравнения* (МПУ). Последние три метода мы рассмотрим в гл. V—VII.

Другая возможность учета многократного рассеяния волн основана на приближенном суммировании рядов теории возмущений (в основном при помощи методов, развитых первоначально в квантовой электродинамике). При таком подходе удастся, в частности, рассмотреть не только слабые, но и сильные флуктуации среды. Однако при этом необходимо, чтобы неоднородности были мелкомасштабными. Элементы теории многократного рассеяния изложены в гл. VIII.

Начнем с простейшей постановки задачи: волновое поле  $u(t, r)$  будем считать *скалярным* и *монохроматическим* ( $u(t, r) = u(r)e^{-i\omega t}$ ), а неоднородности среды — *не меняющимися во времени и покоящимися*<sup>1)</sup>. Хотя при скалярной постановке задачи не охвачена поляризация, она достаточно для ряда общезначимых явлений, таких, как интерференция и дифракция. К поляризационным эффектам в рассеянии электромагнитных волн мы обратимся в § 30.

При указанных выше условиях распространение волны в неоднородной среде описывается уравнением Гельмгольца

$$\Delta u(r) + k_0^2 \varepsilon(r) u(r) = 0, \quad (24.1)$$

<sup>1)</sup> В этой главе мы будем в основном придерживаться монографии [1]. Библиография по вопросам теории однократного рассеяния содержится также в обзоре [2].

где  $k_0 = \omega/c$  — волновое число в невозмущенной среде или в случае электрического поля — в вакууме. Функцию  $\varepsilon(\mathbf{r})$ , описывающую неоднородность среды, мы будем называть (диэлектрической) проницаемостью, имея в виду в основном электромагнитное поле. Для случайно-неоднородной среды проницаемость  $\varepsilon(\mathbf{r})$  можно представить в виде

$$\varepsilon(\mathbf{r}) = \bar{\varepsilon}(\mathbf{r}) + \tilde{\varepsilon}(\mathbf{r}), \quad (24.2)$$

где  $\bar{\varepsilon}$  — среднее (по ансамблю реализаций среды) значение  $\varepsilon$ , а  $\tilde{\varepsilon}$  — флуктуации проницаемости. Уравнение Гельмгольца принимает при этом вид<sup>1)</sup>

$$\Delta u(\mathbf{r}) + k_0^2 [\bar{\varepsilon}(\mathbf{r}) + \tilde{\varepsilon}(\mathbf{r})] u(\mathbf{r}) = 0. \quad (24.3)$$

Общих методов решения даже такого простого волнового уравнения не существует. Наиболее распространенным из приближенных методов является *метод возмущений*: флуктуации  $\tilde{\varepsilon}$  считаются достаточно слабыми, а волновое поле  $u(\mathbf{r})$  ищется в виде ряда по степеням  $\tilde{\varepsilon}(\mathbf{r})$ , или, что то же — по степеням  $\sigma_\varepsilon \ll \bar{\varepsilon}$ . Чтобы построить такой ряд, удобно перейти от дифференциального уравнения (24.3) к эквивалентному интегральному уравнению.

Пусть  $u_0(\mathbf{r})$  — поле первичной волны, удовлетворяющее невозмущенному уравнению Гельмгольца, т. е. уравнению (24.3) при  $\tilde{\varepsilon} = 0$ :

$$\Delta u_0(\mathbf{r}) + k_0^2 \bar{\varepsilon}(\mathbf{r}) u_0(\mathbf{r}) = 0. \quad (24.4)$$

Обозначим через  $G(\mathbf{r}, \mathbf{r}')$  невозмущенную функцию Грина, которая удовлетворяет уравнению для точечного источника

$$\Delta G(\mathbf{r}, \mathbf{r}') + k_0^2 \bar{\varepsilon}(\mathbf{r}) G(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r}, \mathbf{r}'). \quad (24.5)$$

Разумеется, первичное поле  $u_0$  и функция Грина  $G$  удовлетворяют необходимым граничным условиям. Решение неоднородного уравнения

$$\Delta u + k_0^2 \tilde{\varepsilon} u = F(\mathbf{r})$$

выражается через функцию Грина следующим образом:

$$u(\mathbf{r}) = u_0(\mathbf{r}) + \int G(\mathbf{r}, \mathbf{r}') F(\mathbf{r}') d^3 r'. \quad (24.6)$$

<sup>1)</sup> Рассеяние звуковых волн в турбулентной среде описывается более сложным скалярным уравнением, чем (24.3), поскольку распространение сопровождается нелинейным взаимодействием звуковых волн и гидродинамических турбулентных движений. Упрощенное описание достигается путем линеаризации исходных нелинейных уравнений гидродинамики по отношению к слабым звуковым возмущениям. Связанные с этим особенности рассеяния звука рассмотрены в [1].

Записав исходное уравнение (24.3) в форме

$$\Delta u + k_0^2 \bar{\epsilon} u = -k_0^2 \bar{\epsilon} u \equiv F(r) \quad (24.7)$$

и используя (24.6), получаем следующее интегральное уравнение для волнового поля:

$$u(r) = u_0(r) - k_0^2 \int G(r, r') \bar{\epsilon}(r') u(r') d^3 r', \quad (24.8)$$

где интегрирование распространяется, очевидно, на область  $V$ , занятую неоднородностями  $\bar{\epsilon}(r)$ . Уравнение (24.8) эквивалентно исходному дифференциальному уравнению (24.3), но учитывает (через функцию  $G$ ) и все граничные условия задачи.

Ряд теории возмущений строится путем итерирования интегрального уравнения (24.8). Чтобы получить первую итерацию, запишем значение поля в точке  $r = r'$ :

$$u(r') = u_0(r') - k_0^2 \int G(r', r'') \bar{\epsilon}(r'') u(r'') d^3 r'',$$

и подставим это выражение в правую часть (24.8). Это дает

$$u(r) = u_0(r) - k_0^2 \int G(r, r') \bar{\epsilon}(r') u_0(r') d^3 r' + \\ + k_0^4 \int G(r, r') \bar{\epsilon}(r') d^3 r' \int G(r', r'') \bar{\epsilon}(r'') u(r'') d^3 r''. \quad (24.9)$$

Записав значение поля  $u(r)$  в точке  $r = r''$  и подставив его в правую часть (24.9), получим вторую итерацию. Повторяя такую операцию, мы и получим бесконечный ряд теории возмущений:

$$u(r) = u_0(r) - k_0^2 \int G(r, r') \bar{\epsilon}(r') u_0(r') d^3 r' + \\ + k_0^4 \int \int G(r, r') G(r', r'') \bar{\epsilon}(r') \bar{\epsilon}(r'') u_0(r'') d^3 r' d^3 r'' - \\ - k_0^6 \int \int \int G(r, r') G(r', r'') G(r'', r''') \bar{\epsilon}(r') \bar{\epsilon}(r'') \bar{\epsilon}(r''') \times \\ \times u_0(r''') d^3 r' d^3 r'' d^3 r''' + \dots \quad (24.10)$$

В математике этот ряд называется рядом Неймана для интегрального уравнения (24.8), а в физике — борновским разложением<sup>1)</sup>.

Первый член борновского ряда (24.10) — первичное поле  $u_0(r)$ . Второе слагаемое,

$$u_1(r) = -k_0^2 \int G(r, r') \bar{\epsilon}(r') u_0(r') d^3 r', \quad (24.11)$$

<sup>1)</sup> М. Борн впервые применил теорию возмущений в задаче о рассеянии в квантовой механике. Правда, задолго до М. Борна метод возмущений в сходной форме был применен Релем при рассмотрении рассеяния света на прозрачных телах. В оптике и теперь говорят о *релеевском*, а не о *борновском* рассеянии света.

описывает однократно рассеянное поле. Оно порождено непосредственно первичным полем  $u_0(\mathbf{r})$  и линейно относительно возмущений  $\tilde{e}(\mathbf{r})$ . Третье слагаемое в (24.10) можно представить в форме, аналогичной (24.11):

$$u_2(\mathbf{r}) = -k_0^2 \int G(\mathbf{r}, \mathbf{r}') \tilde{e}(\mathbf{r}') u_1(\mathbf{r}') d^3 r'. \quad (24.12)$$

Это—двукратно рассеянное поле, порожденное уже не первичным, а однократно рассеянным полем. Двукратно рассеянное поле в свою очередь возбуждает трехкратно рассеянные волны  $u_3$  и т. д. Таким образом, ряд теории возмущений (24.10) представляет собой разложение *рассеянного поля*  $u_s \equiv u - u_0$  по *кратности рассеяния*:

$$u_s \equiv u - u_0 = u_1 + u_2 + u_3 + \dots \quad (24.13)$$

Из самого способа построения этого ряда видно, что  $n$ -й его член, описывающий  $n$ -кратное рассеяние, содержит под знаком  $n$ -кратного интеграла произведение  $\tilde{e}(\mathbf{r}_1) \dots \tilde{e}(\mathbf{r}_n)$ . Отсюда следует, что для вычисления даже среднего значения поля  $u$  надо знать для  $\tilde{e}$  моменты  $\langle \tilde{e}(\mathbf{r}_1) \dots \tilde{e}(\mathbf{r}_n) \rangle$  любого порядка. При произвольной статистике  $\tilde{e}$  нахождение таких моментов само по себе представляет сложную задачу, но если даже она и разрешима (как, например, в случае нормального распределения), то остается еще открытым вопрос о *методах суммирования усредненных рядов* теории возмущений. В общей постановке этот вопрос будет освещен в гл. VIII. Здесь же, как сказано, мы ограничимся более простой задачей нахождения статистических характеристик поля в приближении *однократного* рассеяния (так называемое *первое борновское* или, чаще, просто *борновское приближение*).

В этом приближении флуктуации  $\tilde{e}$  предполагаются настолько малыми, что в разложении (24.13) можно ограничиться первым членом  $u_1$ . С выражениями такого типа мы уже встречались в § 12 при рассмотрении возбуждения полей случайными источниками. В данном случае в качестве заданных источников  $q(\mathbf{r})$  выступает правая часть уравнения (24.7) с  $u_0$  вместо  $u$ :  $q(\mathbf{r}) = -k_0^2 \tilde{e}(\mathbf{r}) u_0(\mathbf{r})$ . Таким образом, в приближении однократного рассеяния задача о распространении волн в случайно-неоднородных средах (задача типа 2)) сводится к задаче типа 1)—возбуждению полей заданными случайными источниками.

Согласно (24.11) рассеянное поле  $u_s \approx u_1$  является *линейным* функционалом от флуктуаций  $\tilde{e}$ . Поэтому и все моменты поля  $u_1$  *линейно* же выражаются через моменты  $\tilde{e}$  того же порядка. В частности, у однократно рассеянного поля  $u_1$  среднее значение равно нулю, поскольку  $\langle \tilde{e}(\mathbf{r}) \rangle = 0$ , а корреляционная функция  $\Psi_2(\mathbf{r}_1, \mathbf{r}_2) = \langle u_1(\mathbf{r}_1) u_1^*(\mathbf{r}_2) \rangle$  линейно выражается через функцию

корреляции неоднородностей  $\psi_e(\mathbf{r}', \mathbf{r}'')$ :

$$\psi_e(\mathbf{r}_1, \mathbf{r}_2) = k_0^2 \int \int G(\mathbf{r}_1, \mathbf{r}') G^*(\mathbf{r}_2, \mathbf{r}'') u_0(\mathbf{r}') u_0^*(\mathbf{r}'') \times \\ \times \psi_e(\mathbf{r}', \mathbf{r}'') d^3\mathbf{r}' d^3\mathbf{r}''. \quad (24.14)$$

Выражение (24.14) и аналогичные квадратуры для высших моментов рассеянного поля  $u_1$  в принципе дают *полное* статистическое решение задачи в рассматриваемом борновском приближении. Однако этот математический результат еще нуждается в физическом истолковании.

### § 25. Средняя интенсивность рассеянного поля

Для того чтобы лучше уяснить основные закономерности рассеяния, сделаем ряд допущений, которые упрощают анализ, но вместе с тем сохраняют общность, достаточную для многих приложений теории. Допущения сводятся к следующему.

а) Среда в среднем однородна, т. е.  $\epsilon = \text{const}$ .

б) Первичное поле  $u_0(\mathbf{r})$  представляет собой ненаправленную сферическую волну с центром в точке  $\mathbf{r}_0$ :

$$u_0(\mathbf{r}) = \frac{A}{|\mathbf{r} - \mathbf{r}_0|} e^{ik|\mathbf{r} - \mathbf{r}_0|}, \quad (25.1)$$

где  $k$  — волновое число в однородной среде:  $k = k_0 \sqrt{\epsilon} = \frac{\omega}{c} \sqrt{\epsilon}$ .

в) Функция Грина  $G(\mathbf{r}, \mathbf{r}')$  описывает поле точечного источника в неограниченной однородной среде:

$$G(\mathbf{r}, \mathbf{r}') = -\frac{1}{4\pi} \frac{e^{ik|\mathbf{r} - \mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|}. \quad (25.2)$$

г) Поле флуктуаций  $\tilde{\epsilon}$  статистически квазиоднородно, т. е. корреляционная функция  $\psi_e$  имеет вид (§ 5)

$$\psi_e(\mathbf{r}', \mathbf{r}'') \equiv \langle \tilde{\epsilon}(\mathbf{r}') \tilde{\epsilon}(\mathbf{r}'') \rangle = \psi_e(\boldsymbol{\rho}, \mathbf{R}), \quad (25.3)$$

где  $\boldsymbol{\rho} = \mathbf{r}' - \mathbf{r}''$ ,  $\mathbf{R} = (\mathbf{r}' + \mathbf{r}'')/2$  и зависимость  $\psi_e(\boldsymbol{\rho}, \mathbf{R})$  от  $\mathbf{R}$  «медленная», т. е. масштаб  $L_e$  изменения  $\psi_e$  по аргументу  $\mathbf{R}$  существенно больше, чем характерный масштаб  $l_e$  (радиус корреляции  $\tilde{\epsilon}$ ) по разности аргументу  $\boldsymbol{\rho}$ . Спектральная плотность таких флуктуаций, определяемая выражением

$$\Phi_{ee}(\mathbf{x}, \mathbf{R}) = \frac{1}{8\pi^3} \int \psi_e(\boldsymbol{\rho}, \mathbf{R}) \exp(-i\mathbf{x}\boldsymbol{\rho}) d^3\boldsymbol{\rho}, \quad (25.4)$$

— медленная функция  $\mathbf{R}$ , т. е. она мало меняется на расстояниях порядка  $l_e$ .

д) Наконец, будем считать, что случайные неоднородности заполняют *ограниченный* объем  $V$  и в этом объеме содержится

много неоднородностей. Последнее условие можно записать в виде неравенства

$$V \sim L^3 \gg l_e^3, \quad \text{или} \quad L \sim V^{1/3} \gg l_e, \quad (25.5)$$

где  $L$  — поперечный размер области, занятой неоднородностями. Предположение о конечности рассеивающего объема необходимо для обеспечения малости однократно рассеянного поля, тогда как неравенство (25.5) принято лишь для упрощения расчетов.

Конечность рассеивающего объема удобно учитывать при помощи обрезающей функции

$$M(\mathbf{r}) = \begin{cases} 1 & \text{внутри } V, \\ 0 & \text{вне } V. \end{cases} \quad (25.6)$$

Введя ее под знак интеграла в (24.11), можно распространить интегрирование на все пространство. Для первичной сферической волны (25.1) и функции Грина свободного пространства (25.2) однократно рассеянное поле (24.11) запишется в виде

$$u_1(\mathbf{r}) = \frac{k_0^3 A}{4\pi} \int_{-\infty}^{\infty} \frac{\exp[ik(|\mathbf{r}-\mathbf{r}'| + |\mathbf{r}'-\mathbf{r}_0|)]}{|\mathbf{r}-\mathbf{r}'||\mathbf{r}'-\mathbf{r}_0|} M(\mathbf{r}') \tilde{\epsilon}(\mathbf{r}') d^3 r'. \quad (25.7)$$

В дальнейшем мы обратимся к некоторым более общим постановкам задачи (учет временных изменений  $\tilde{\epsilon}$ , расчет поляризационных характеристик электромагнитного поля и др.). Вместе с тем мы будем иногда вводить частные допущения (плоская первичная волна, статистически однородные флуктуации и т. д.). Здесь же, исходя из (25.7), мы получим выражения для функции корреляции и, в частности, для средней интенсивности рассеянного поля.

Согласно (25.7) пространственная функция корреляции поля равна

$$\begin{aligned} \psi_n(\mathbf{r}_1, \mathbf{r}_2) &= \frac{k_0^4 |A|^2}{(4\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} M(\mathbf{r}') M(\mathbf{r}'') \psi_e(\mathbf{p}, \mathbf{R}) \times \\ &\times \frac{\exp[ik(|\mathbf{r}_1-\mathbf{r}'| + |\mathbf{r}_2-\mathbf{r}''| + |\mathbf{r}'-\mathbf{r}_0| + |\mathbf{r}''-\mathbf{r}_0|)]}{|\mathbf{r}_1-\mathbf{r}'||\mathbf{r}_2-\mathbf{r}''||\mathbf{r}'-\mathbf{r}_0||\mathbf{r}''-\mathbf{r}_0|} d^3 r' d^3 r''. \end{aligned} \quad (25.8)$$

От выражения (24.14) эта формула отличается только тем, что в ней конкретизированы вид первичного поля и вид функции Грина.

Перейдем в (25.8) к переменным интегрирования  $\mathbf{p} = \mathbf{r}' - \mathbf{r}''$ ,  $\mathbf{R} = (\mathbf{r}' + \mathbf{r}'')/2$ , откуда  $\mathbf{r}' = \mathbf{R} + \mathbf{p}/2$ ,  $\mathbf{r}'' = \mathbf{R} - \mathbf{p}/2$ , причем  $d^3 r' d^3 r'' = d^3 \mathbf{p} d^3 \mathbf{R}$ .

Формула (25.8) принимает вид

$$\begin{aligned} \Psi_{\mathbf{r}}(\mathbf{r}_1, \mathbf{r}_2) = \\ = \frac{k_0^4 |A|^2}{(4\pi)^3} \iint_{-\infty}^{\infty} \frac{e^{i(\Psi_1 + \Psi_2)} M\left(\mathbf{R} + \frac{\boldsymbol{\rho}}{2}\right) M\left(\mathbf{R} - \frac{\boldsymbol{\rho}}{2}\right) \Psi_{\mathbf{e}}(\boldsymbol{\rho}, \mathbf{R}) d^3\rho d^3R}{\left|\mathbf{r}_1 - \mathbf{R} - \frac{\boldsymbol{\rho}}{2}\right| \left|\mathbf{r}_2 - \mathbf{R} + \frac{\boldsymbol{\rho}}{2}\right| \left|\mathbf{R} + \frac{\boldsymbol{\rho}}{2} - \mathbf{r}_0\right| \left|\mathbf{R} - \frac{\boldsymbol{\rho}}{2} - \mathbf{r}_0\right|}, \end{aligned} \quad (25.9)$$

где

$$\begin{aligned} \Psi_1 = k(|\mathbf{r}_1 - \mathbf{r}'| - |\mathbf{r}_2 - \mathbf{r}''|) = k\left(\left|\mathbf{r}_1 - \mathbf{R} - \frac{\boldsymbol{\rho}}{2}\right| - \left|\mathbf{r}_2 - \mathbf{R} + \frac{\boldsymbol{\rho}}{2}\right|\right), \\ \Psi_2 = k(|\mathbf{r}' - \mathbf{r}_0| - |\mathbf{r}'' - \mathbf{r}_0|) = k\left(\left|\mathbf{R} - \mathbf{r}_0 + \frac{\boldsymbol{\rho}}{2}\right| - \left|\mathbf{R} - \mathbf{r}_0 - \frac{\boldsymbol{\rho}}{2}\right|\right). \end{aligned} \quad (25.10)$$

Все последующие упрощения формулы (25.9) основаны на том, что функция корреляции неоднородной среды  $\Psi_{\mathbf{e}}(\boldsymbol{\rho}, \mathbf{R})$  становится при  $\rho \gg l_e$  очень малой. Мы можем поэтому воспользоваться разложением модуля вектора  $|\mathbf{r} - \boldsymbol{\rho}|$  в ряд Тейлора по степеням малого отношения  $\rho/r$ :

$$|\mathbf{r} - \boldsymbol{\rho}| = \sqrt{(\mathbf{r} - \boldsymbol{\rho})^2} \approx r - \mathbf{n}\boldsymbol{\rho} + \frac{1}{2r} \rho_{\perp}^2 + O\left(\frac{\rho^3}{r^2}\right), \quad (25.11)$$

где  $\mathbf{n} = \mathbf{r}/r$ , а  $\rho_{\perp}$  — перпендикулярная к  $\mathbf{n}$  составляющая вектора  $\boldsymbol{\rho}$ :  $\rho_{\perp}^2 = \rho^2 - (\mathbf{n}\boldsymbol{\rho})^2$ .

Применим разложение (25.11) к каждому из слагаемых, входящих в  $\Psi_1$  и  $\Psi_2$ . Для  $\Psi_1$  получаем

$$\begin{aligned} \Psi_1 = k \left[ s - \frac{1}{2} (\mathbf{n}_{s1} + \mathbf{n}_{s2}) \boldsymbol{\rho} \cdot \frac{\boldsymbol{\rho}_{\perp 1}}{8|\mathbf{r}_1 - \mathbf{R}|} - \frac{\boldsymbol{\rho}_{\perp 2}}{8|\mathbf{r}_2 - \mathbf{R}|} + \right. \\ \left. + O\left(\frac{\rho^3}{|\mathbf{r}_{1,2} - \mathbf{R}|^2}\right) \right]. \end{aligned} \quad (25.12)$$

Здесь

$$s = |\mathbf{r}_1 - \mathbf{R}| - |\mathbf{r}_2 - \mathbf{R}| \quad (25.13)$$

— разность хода от «точки рассеяния»  $\mathbf{R}^1$  до точек наблюдения  $\mathbf{r}_1$  и  $\mathbf{r}_2$ ;  $\mathbf{n}_{s1}$  и  $\mathbf{n}_{s2}$  — единичные векторы, направленные на точки  $\mathbf{r}_1$  и  $\mathbf{r}_2$  (рис. 23):

$$\mathbf{n}_{s1} = \frac{\mathbf{r}_1 - \mathbf{R}}{|\mathbf{r}_1 - \mathbf{R}|}, \quad \mathbf{n}_{s2} = \frac{\mathbf{r}_2 - \mathbf{R}}{|\mathbf{r}_2 - \mathbf{R}|}, \quad (25.14)$$

а  $\rho_{\perp 1}$  и  $\rho_{\perp 2}$  — перпендикулярные к  $\mathbf{n}_{s1}$  и  $\mathbf{n}_{s2}$  составляющие вектора  $\boldsymbol{\rho}$ .

<sup>1)</sup> Точнее, от «центра тяжести», т. е. от точки, лежащей посередине между точками рассеяния  $\mathbf{r}'$  и  $\mathbf{r}''$ , раздвинутыми на малое расстояние  $\rho \ll l_e$ .

Выясним условия, при которых в (25.12) можно пренебречь квадратичными и кубическими по  $\rho$  членами. Фактическая область

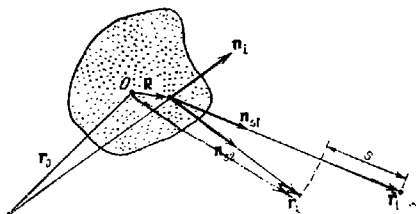


Рис. 23.

интегрирования по  $\rho$  выделена неравенством  $\rho \leq l_e$ . Требуя малости квадратичных членов в (25.12) по сравнению с  $\pi$  при  $\rho \sim l_e$ , получаем неравенство

$$\frac{k}{8} \left( \frac{\rho_{\perp 1}^3}{|r_1 - R|} - \frac{\rho_{\perp 2}^3}{|r_2 - R|} \right) \Big|_{\rho \sim l_e} \ll \pi, \quad (25.15)$$

которое заведомо будет выполнено, если оно справедливо для каждого из слагаемых:

$$\frac{k \rho_{\perp 1,2}^3}{8 |r_{1,2} - R|} \Big|_{\rho \sim l_e} \ll \pi.$$

Смысл же последних неравенств заключается в том, что точки наблюдения  $r_1$  и  $r_2$  должны находиться во *фраунгоферовой* зоне по отношению к неоднородностям размера  $l_e$ :

$$|r_{1,2} - R| \gg \frac{k l_e^2}{8\pi} \sim \frac{l_e^2}{\lambda}. \quad (25.16)$$

Ясно, однако, что неравенство (25.16) следует рассматривать как *достаточное* условие пренебрежения высшими степенями  $\rho$ , поскольку в (25.15) входит *разность* двух примерно одинаковых по величине слагаемых. Если точки наблюдения совмещены ( $r_1 = r_2$ ), то левая часть (25.15) обращается в нуль и квадратичные члены в разложении (25.12) просто отсутствуют. Вместо (25.15) следует потребовать в этом случае малости (по сравнению с  $\pi$ ) уже не квадратичных, а кубических членов:

$$\frac{k \rho^3}{|r_{1,2} - R|^2} \Big|_{\rho \sim l_e} \ll \pi.$$

Это неравенство налагает на расстояния  $|r_{1,2} - R|$  ограничение

$$|r_{1,2} - R| \gg \sqrt{k l_e^3} \approx \sqrt{l_e^3 / \lambda}. \quad (25.17)$$



В случае крупных неоднородностей ( $l_e \gg \lambda$ ) условие (25.17) значительно слабее условия (25.16), согласно которому точки наблюдения должны находиться во френгоферовой зоне отдельной неоднородности (размер  $\sim l_e$ ). Напротив, при  $l_e \ll \lambda$  неравенство (25.17) сильнее условия (25.16). Для неоднородностей, размеры которых сравнимы с длиной волны, оба условия (25.16) и (25.17) равносильны.

Что касается слагаемого  $\Psi_2$ , то разложение этой величины содержит только нечетные степени  $\rho$ . Поэтому при выполнении аналогичного (25.17) условия

$$|R - r_0| \gg \sqrt{l_e^2/\lambda}, \quad (25.18)$$

позволяющего отбросить кубические члены разложения, имеем

$$\Psi_2 \approx kn_i \rho,$$

где

$$n_i(R) = \frac{R - r_0}{|R - r_0|} \quad (25.19)$$

— единичный вектор в направлении распространения первичной волны в точке  $R$  (рис. 23). В результате

$$\Psi = \Psi_1 + \Psi_2 \approx k \left[ s^{-1/2} (n_{s1} + n_{s2}) \rho - n_i \rho \right]. \quad (25.20)$$

В знаменателе подынтегрального выражения (25.9) и в произведении  $M(R + \rho/2) M(R - \rho/2)$  мы просто пренебрежем  $\rho$ . Для такого упрощения знаменателя достаточно, чтобы обе точки наблюдения  $r_1$  и  $r_2$  и точка излучения  $r_0$  отстояли от точки рассеяния  $R$  не меньше, чем на  $l_e$ . Замена же произведения  $M(R + \rho/2) \times M(R - \rho/2)$  на  $M^2(R)$  или, что то же, на  $M(R)$  пригодна во всем объеме  $V$ , за исключением приграничного слоя толщины порядка  $l_e$ . Совершаемая при этой замене относительная ошибка — порядка отношения объема пограничного слоя  $L^3 l_e$  к полному объему  $V \sim L^3$ , т. е. порядка  $l_e/L \ll 1$  (см. (25.5)).

При сделанных допущениях формула (25.9) принимает вид

$$\Psi_2(r_1, r_2) = \frac{k_0^4 |A|^2}{(4\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{M(R) \exp(iks - iQ\rho) \psi_e(\rho, R)}{|r_1 - R| |r_2 - R| |R - r_0|^2} d^3\rho d^3R, \quad (25.21)$$

где введено обозначение

$$Q(R) = k \left\{ \frac{1}{2} [n_{s1}(R) + n_{s2}(R)] - n_i(R) \right\}. \quad (25.22)$$

Согласно (25.4) интеграл по  $\rho$  можно выразить через спектральную плотность квазигомогенных флуктуаций:

$$\int_{-\infty}^{\infty} \psi_e(\rho, R) \exp(-iQ\rho) d^3\rho = 8\pi^3 \Phi_e(Q, R), \quad (25.23)$$

так что окончательно получаем

$$\psi_{\pi}(\mathbf{r}_1, \mathbf{r}_2) = \frac{\pi k_0^4 |A|^2}{2} \int_{-\infty}^{\infty} \frac{M(\mathbf{R}) \Phi_{\pi}(\mathbf{Q}, \mathbf{R}) e^{ik_s R}}{|\mathbf{r}_1 - \mathbf{R}| |\mathbf{r}_2 - \mathbf{R}| |\mathbf{R} - \mathbf{r}_0|^2} d^3 R. \quad (25.24)$$

Анализ этой функции корреляции мы отложим до § 27, а здесь рассмотрим только пространственное распределение *средней интенсивности*

$$\bar{I}_1(\mathbf{r}) = \langle |u_1(\mathbf{r})|^2 \rangle = \psi_{\pi}(\mathbf{r}, \mathbf{r}). \quad (25.25)$$

При совмещении точек наблюдения ( $\mathbf{r}_1 = \mathbf{r}_2 = \mathbf{r}$ , так что  $s = 0$ ) вектор  $\mathbf{Q}$  (25.22) переходит в так называемый *вектор рассеяния*

$$\mathbf{q}(\mathbf{R}) = k[\mathbf{n}_s(\mathbf{R}) - \mathbf{n}_i(\mathbf{R})]. \quad (25.26)$$

В результате средняя интенсивность будет

$$\bar{I}_1(\mathbf{r}) = \frac{\pi k_0^4 |A|^2}{2} \int_{-\infty}^{\infty} \frac{M(\mathbf{R}) \Phi_{\pi}(\mathbf{q}, \mathbf{R})}{|\mathbf{r} - \mathbf{R}|^2 |\mathbf{R} - \mathbf{r}_0|^2} d^3 R. \quad (25.27)$$

Как видно отсюда, волны, рассеянные отдельными элементами объема  $d^3 R$ , складываются *некогерентно*, т. е. складываются их интенсивности: выражение под интегралом (конечно, вместе с коэффициентом, вынесенным в (25.27) из-под интеграла) представляет собой среднюю интенсивность  $d\bar{I}_1$  поля, рассеянного элементом объема  $d^3 R$ .

Формулу (25.27), выведенную для сферической первичной волны (25.1), нетрудно распространить на волны другой формы. Величина  $I_0(\mathbf{R}) = |A|^2 / |\mathbf{R} - \mathbf{r}_0|^2$  — это интенсивность сферической волны в точке  $\mathbf{R}$ . Записав (25.27) в виде

$$\bar{I}_1(\mathbf{r}) = \frac{\pi k_0^4}{2} \int_{-\infty}^{\infty} \frac{M(\mathbf{R}) I_0(\mathbf{R}) \Phi_{\pi}(\mathbf{q}, \mathbf{R})}{|\mathbf{r} - \mathbf{R}|^2} d^3 R, \quad (25.28)$$

мы получаем выражение, справедливое для *любого* первичного поля, если его структура в пределах одной неоднородности размера  $l_e$  практически не отличается от структуры сферической волны. Последнее означает, что в пределах отдельной неоднородности амплитуду поля и радиус кривизны  $R_{кр}$  волнового фронта можно считать постоянными, причем радиус  $R_{кр}$  должен удовлетворять еще неравенствам вида (25.18), т. е.  $R_{кр} \gg \sqrt{l_e^3 / \lambda}$ .

Таким образом, мы можем применить выражение (25.28) и к *плоской* первичной волне

$$u_i = A_0 e^{ik n_i r}, \quad (25.29)$$

для которой  $I_0 = |A_0|^2 = \text{const}$  (в отличие от (25.19), единичный вектор  $\mathbf{n}_i$ , указывающий направление распространения первичной волны, здесь постоянен во всем пространстве), и к *направленной сферической волне*

$$u_i = A f_0(\mathbf{n}_i) \frac{e^{ik|\mathbf{r}-\mathbf{r}_0|}}{|\mathbf{r}-\mathbf{r}_0|} \quad (25.30)$$

( $f_0(\mathbf{n}_i)$  — диаграмма направленности излучателя), для которой

$$I_0(\mathbf{R}) = \frac{|A|^2 |f_0(\mathbf{n}_i)|^2}{|\mathbf{R}-\mathbf{r}_0|^2}. \quad (25.31)$$

Наиболее существенная особенность формул (25.27) и (25.28) заключается в том, что они отражают *селективный* характер рассеяния. Пусть  $\theta$  — угол между волновыми векторами падающей ( $\mathbf{k}_i$ ) и рассеянной ( $\mathbf{k}_s$ ) волн, называемый *углом рассеяния*. Очевидно,

$$q = k|\mathbf{n}_s - \mathbf{n}_i| = 2k \sin \frac{1}{2}\theta. \quad (25.32)$$

Этому значению  $q$  отвечает пространственная гармоника возмущений  $e^{iq\mathbf{r}}$  с длиной волны

$$\Lambda_q = \frac{2\pi}{q} = \frac{\lambda}{2 \sin(\theta/2)}, \quad (25.33)$$

а направление нормали к фронту параллельно вектору  $\mathbf{q}$ , т. е. параллельно  $(\mathbf{n}_s - \mathbf{n}_i)$  (рис. 24). Следовательно, равенство (25.33) — это условие Вульфа — Брегга, определяющее пространственный период той гармоники, на которой волна дифрагирует под углом  $\theta$ .

Разумеется, пространственный спектр флуктуаций  $\hat{\varepsilon}(\mathbf{r})$  содержит бесконечный набор таких гармоник (объемных «дифракционных решеток») со всевозможными периодами и ориентациями. Селективность рассеяния заключается в том, что в заданном направлении  $\mathbf{n}_s$  дают вклад волны, рассеянные лишь на выделенной решетке — с пространственным периодом (25.33) и с ориентацией, отвечающей закону отражения (в точке  $\mathbf{R}$ ) первичной волны из точки  $\mathbf{r}_0$  в точку  $\mathbf{r}$ . Интенсивность этой гармоники пропорциональна спектральной плотности  $\Phi_\varepsilon(\mathbf{q})$ , которая и фигурирует в (25.27) и (25.28).

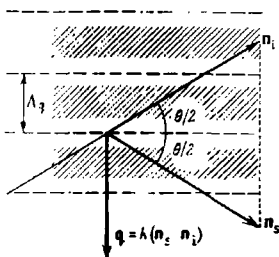


Рис. 24.

При *обратном* рассеянии, когда  $k_s = -k_i$ , так что вектор рассеяния равен  $q = -2k_i$ , а  $\theta = \pi$ , рассеяние обусловлено гармоникой с периодом  $\Lambda_q = \lambda/2$ , т. е. с периодом, равным половине длины волны падающего излучения. Обратное рассеяние наблюдается в очень многих случаях, например в радио- и гидролокации, когда источник и приемник излучения расположены в одной точке.

Рассеяние под произвольным углом часто используется в оптике при изучении неоднородностей в прозрачных материалах.

При уменьшении угла рассеяния  $\theta$  период  $\Lambda_q$  увеличивается. Значения  $\Lambda_q$  и  $\lambda$  равны друг другу при  $\theta = \pi/3$ . При дальнейшем уменьшении  $\theta$  пространственный период рассеивающей решетки  $\Lambda_q$  превышает  $\lambda$  и для рассеяния *вперед* ( $\theta \rightarrow 0$ ) величина  $q$  обращается в

нуль, а  $\Lambda_q$  — в бесконечность. Это означает, что рассеяние *вперед* обусловлено наиболее крупными объемными возмущениями с пространственными масштабами  $\Lambda_q \gg \lambda$ .

Формула (25.28) упрощается для столь малого рассеивающего объема, что в его пределах величины  $\Phi_e(q, R)$ ,  $I_0(R)$  и  $1/|r - R|^2$  практически постоянны. Вынося их из-под знака интеграла со значениями, соответствующими, скажем, центру  $O$  рассеивающего объема ( $R = 0$ ), получаем

$$\bar{I}_1(r) = \frac{\pi k_0^4}{2} \frac{I_0(0) \Phi_e(q_0, 0)}{r^2} V. \quad (25.34)$$

Здесь  $q_0 = q(0) = k(n_{s0} - n_{i0})$  — вектор рассеяния, отвечающий центру рассеивающего объема, а  $n_{s0} = r/r$  и  $n_{i0} = r_0/r_0$  — единичные векторы, направленные из  $O$  (рис. 25).

Замена множителя  $1/|r - R|^2$  на  $1/r^2$  возможна при очевидном условии, что точка наблюдения удалена на расстояние, значительно превышающее поперечный размер  $L$  рассеивающего объема:

$$r \gg L. \quad (25.35)$$

Если падающая волна сферическая, то переход от точного выражения для интенсивности  $I_0(R) = |A|^2/|R - r_0|^2$  к приближенному значению  $I_0(0) = |A|^2/r_0^2$  требует выполнения такого же неравенства и для расстояния до источника:  $r_0 \gg L$ .

Что касается замены  $\Phi_e(q, R)$  постоянным значением  $\Phi_e(q_0, 0)$ , то для этого необходима, во-первых, статистическая *однородность*

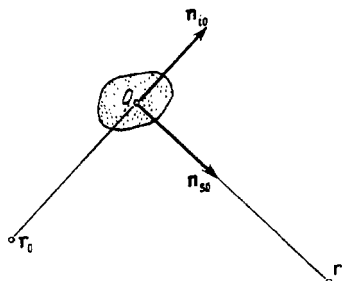


Рис. 25.

флуктуаций в пределах  $V$  (масштаб  $l_e$  должен быть больше поперечника  $L$ ) и, во-вторых, малость изменения (в пределах объема  $V$ ) вектора  $\mathbf{q}$  по сравнению с характерным масштабом изменения спектра  $\Phi_e(\mathbf{q}, 0)$ , равным  $1/l_e$ :

$$\Delta q \ll 1/l_e. \quad (25.36)$$

По порядку величины  $\Delta q \sim k \Delta n_s$  или  $k \Delta n_l$  ( $\Delta n_s$  и  $\Delta n_l$  — вариации направлений падающей и рассеянных волн в пределах  $V$ ). Очевидно,  $\Delta n_s \sim L/r$ , так что неравенство (25.36) сводится к условию

$$r \gg k L l_e \sim L l_e / \lambda. \quad (25.37)$$

Поскольку ограничения для расстояния  $r_0$  до источника сферической волны формулируются аналогичным образом, мы будем далее говорить только о неравенствах, относящихся к расстоянию  $r$  до точки наблюдения.

В случае мелкомасштабных неоднородностей ( $l_e \ll \lambda$ ) из двух неравенств (25.35) и (25.37) более жестким, очевидно, является первое, тогда как для крупномасштабных неоднородностей ( $l_e \gg \lambda$ ) минимальное расстояние до точки наблюдения ограничено условием (25.37). Обозначив через  $r_{\min}$  меньшую из величин  $L$  и  $L l_e / \lambda$ , можно записать оба неравенства (25.35) и (25.37) в виде

$$r \gg r_{\min} = \min \{ L, L l_e / \lambda \}. \quad (25.38)$$

Обычно упрощенное выражение (25.34) выводят в предположении, что точка наблюдения находится во френгоферовой зоне всего объема  $V$ . Для этого в выражении (25.7) (или в аналогичном ему выражении для плоской первичной волны) разлагают модуль разности  $|\mathbf{r} - \mathbf{r}'|$  в ряд по степеням  $\mathbf{r}'$  и пренебрегают квадратичным слагаемым  $k r_1^2 / 2r$  непосредственно в функции Грина (25.2). Это и приводит к френгоферову приближению

$$u_1(\mathbf{r}) = \frac{k_0^2 A e^{i k r}}{4 \pi r r_0} \int_{-\infty}^{\infty} e^{-i \mathbf{q} \cdot \mathbf{r}'} M(\mathbf{r}') \tilde{\mathbf{e}}(\mathbf{r}') d^3 r'. \quad (25.39)$$

Выражение для средней интенсивности, вычисляемое при помощи (25.39), значительно проще, чем в общем случае, а именно совпадает с (25.34). Однако условие

$$\frac{k r'^2}{2r} \ll \frac{k L^2}{2r} \ll \pi, \text{ или } r \gg \frac{L^2}{\lambda}, \quad (25.40)$$

которое необходимо для френгоферова приближения (25.39), гораздо жестче, чем условие (25.38), достаточное для применимости (25.34).

Во многих прикладных задачах измеряется не само рассеянное поле  $u_1(\mathbf{r})$ , а сигнал на выходе приемной антенны. Если

центр приемной апертуры расположен в точке  $\mathbf{r}$ , то выходной сигнал  $v$  представляет собой поле  $u_1(\mathbf{r} + \boldsymbol{\rho})$ , проинтегрированное по апертуре  $\Sigma$  с весовой функцией  $b(\boldsymbol{\rho})$ , которая описывает распределение тока в антенне в передающем режиме:

$$v = \widetilde{u}_1(\mathbf{r}) = \int_{\Sigma} u_1(\mathbf{r} + \boldsymbol{\rho}) b(\boldsymbol{\rho}) d^2\rho. \quad (25.41)$$

Интегрирование по апертуре (волнистая черта), разумеется, не равносильно статистическому усреднению.

Подставим в (25.41) рассеянное поле (25.7), не конкретизируя пока вида первичной волны  $u_0(\mathbf{r})$ :

$$v = \int_{\Sigma} b(\boldsymbol{\rho}) d^2\rho \int \frac{k_0^3}{4\pi} \frac{\exp(ik|\mathbf{r} - \boldsymbol{\rho} - \mathbf{r}'|)}{|\mathbf{r} + \boldsymbol{\rho} - \mathbf{r}'|} M(\mathbf{r}') \tilde{e}(\mathbf{r}') u_0(\mathbf{r}') d^3r'. \quad (25.42)$$

Величину  $|\mathbf{r} + \boldsymbol{\rho} - \mathbf{r}'|$  разложим в ряд Тейлора по  $\boldsymbol{\rho}$ :

$$|\mathbf{r} + \boldsymbol{\rho} - \mathbf{r}'| = |\mathbf{r} - \mathbf{r}'| + n'_s \boldsymbol{\rho} + \frac{1}{2|\mathbf{r} - \mathbf{r}'|} [\boldsymbol{\rho}^2 - (n'_s \boldsymbol{\rho})^2] + \dots, \quad (25.43)$$

где  $n'_s \equiv \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|}$  — единичный вектор, направленный из точки рассеяния  $\mathbf{r}'$  в центр приемной апертуры  $\mathbf{r}$  (рис. 26).

Пусть  $d$  — наибольший размер приемной антенны, так что в пределах ее апертуры  $\rho \leq d$ , и пусть выполнено условие  $kd^2 \ll |\mathbf{r} - \mathbf{r}'|$  (точка  $\mathbf{r}'$  лежит во френгоферовой зоне антенны, где уже сформировалась диаграмма направленности). Тогда в показателе экспоненты в (25.42) можно ограничиться двумя первыми членами разложения (25.43), а в знаменателе подынтегрального выражения — первым членом этого разложения. Выполнив интегрирование по  $\boldsymbol{\rho}$ , находим

$$v = \frac{k_0^3}{4\pi} \int f_1(n'_s) \frac{e^{ik|\mathbf{r} - \mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} M(\mathbf{r}') u_0(\mathbf{r}') \tilde{e}(\mathbf{r}') d^3r', \quad (25.44)$$

где

$$f_1(n'_s) = \int_{\Sigma} b(\boldsymbol{\rho}) \exp(ikn'_s \boldsymbol{\rho}) d^2\rho \quad (25.45)$$

— диаграмма направленности приемной антенны. Максимальное значение  $|f_1(n'_s)|$  удобно нормировать к единице. Если весовая функция  $b(\boldsymbol{\rho})$  симметрична относительно центра  $\boldsymbol{\rho} = 0$ , то максимум  $|f_1(n'_s)|$  достигается при  $n'_s \boldsymbol{\rho} = 0$ , т. е. при ориентации вектора  $n'_s$  по нормали к плоскости антенны.

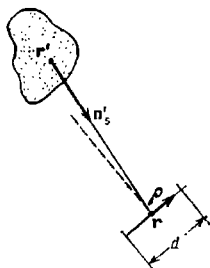


Рис. 26.

От исходной формулы (25.7) выражение (25.44) отличается тем, что под интегралом появился диаграммный множитель  $f_1(n'_s)$ . При малых по сравнению с длиной волны размерах приемной антенны ( $kd \ll 1$ ) этот множитель не зависит от направления  $n'_s$ , и мы возвращаемся тогда к формуле (25.7).

Взяв в качестве первичной направленную сферическую волну (25.30), получаем для  $v$  выражение

$$v = \frac{k_0^2 A}{4\pi} \int_{-\infty}^{\infty} M(r') \frac{e^{ik(|r-r'|+|r'-r_0|)}}{|r-r'| |r'-r_0|} f_0(n'_i) f_1(n'_s) \tilde{e}(r') d^3 r', \quad (25.46)$$

которое, если ввести обозначение

$$\mathcal{D}(r') = \frac{k_0^2 A}{4\pi} M(r') \frac{f_0(n'_i) f_1(n'_s)}{|r-r'| |r'-r_0|}, \quad (25.47)$$

можно записать в виде

$$v = \int_{-\infty}^{\infty} \mathcal{D}(r') e^{ik(|r-r'|+|r'-r_0|)} \tilde{e}(r') d^3 r'. \quad (25.48)$$

Заметим, что в рассматриваемом случае, когда излучение и прием осуществляются направленными антеннами, область интегрирования в (25.46) или (25.48) остается конечной и тогда, когда неоднородности  $\tilde{e}(r)$  заполняют все пространство, т. е.  $M(r')$  всюду равно единице. Специальное ограничение величины занятого неоднородностями объема теперь излишне, так как произведение диаграммных множителей спадает практически до нуля вне области пересечения центральных (главных) лепестков обеих диаграмм (рис. 27).

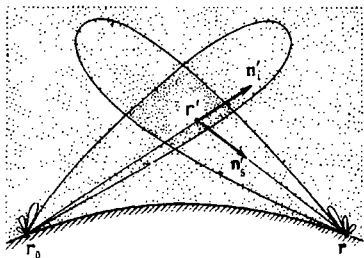


Рис. 27.

Располагая выражением (25.46), нетрудно вычислить среднюю интенсивность сигнала на выходе антенны  $\bar{I}_v \equiv \langle |v|^2 \rangle$ . В отличие от (25.27), под знак интеграла теперь войдет множитель  $|f_0(n_i) f_1(n_s)|^2$ :

$$\bar{I}_v = \frac{\pi k_0^4 |A|^2}{2} \int_{-\infty}^{\infty} \frac{M(R) \Phi_e(q, R)}{|r-R|^2 |R-r_0|^2} |f_0(n_i)|^2 |f_1(n_s)|^2 d^3 R. \quad (25.49)$$

При использовании остронаправленных передающих и приемных антенн интенсивность рассеянного поля удобно характеризовать величиной *эффективного рассеивающего объема*. Обозначим через  $\mathbf{q}_0$  вектор рассеяния  $k(\mathbf{n}_{i0} - \mathbf{n}_{s0})$ ; отвечающий точке пересечения максимумов диаграмм направленности, которую естественно принять за начало координат, и запишем (25.49) в форме, аналогичной выражению (25.34) ( $I_0 = |A|^2/r_0^2$ ):

$$\bar{I}_v = \frac{\pi k_0^4 \Phi_e(\mathbf{q}_0, 0) I_0}{2r^2} V_{\text{эфф}}. \quad (25.50)$$

Множитель

$$V_{\text{эфф}} = \int M(\mathbf{R}) \left[ \frac{rr_0}{|\mathbf{r} - \mathbf{R}| |\mathbf{R} - \mathbf{r}_0|} \right]^2 |f_0(\mathbf{n}_i) f_1(\mathbf{n}_s)|^2 \frac{\Phi_e(\mathbf{q}, \mathbf{R})}{\Phi_e(\mathbf{q}_0, 0)} d^3 R, \quad (25.51)$$

определенный из сравнения (25.49) с (25.34), и представляет собой эффективный рассеивающий объем.

Величина  $V_{\text{эфф}}$ , как следует из (25.51), зависит от многих факторов: от расстояний до источника и приемника, от формы диаграмм направленности антенн и от вида спектра  $\Phi_e(\mathbf{q}, \mathbf{R})$ . При использовании остронаправленных антенн, если область пересечения диаграмм целиком лежит в рассеивающем объеме, величины  $M(\mathbf{R})$  и  $rr_0/|\mathbf{R} - \mathbf{r}_0| |\mathbf{r} - \mathbf{R}|$  можно с хорошим приближением принять равными единице, и тогда

$$V_{\text{эфф}} = \int_{-\infty}^{\infty} |f_0(\mathbf{n}_i) f_1(\mathbf{n}_s)|^2 \frac{\Phi_e(\mathbf{q}, \mathbf{R})}{\Phi_e(\mathbf{q}_0, 0)} d^3 R. \quad (25.52)$$

Эта формула часто используется при расчетах интенсивности радиоволн, рассеянных в тропосфере и нижней ионосфере [1].

## § 26. Эффективный поперечник рассеяния.

Границы применимости приближения  
однократного рассеяния

Для скалярного монохроматического поля  $u(\mathbf{r})$ , удовлетворяющего уравнению Гельмгольца (24.1), в случае прозрачной среды, т. е. вещественной проницаемости  $\varepsilon(\mathbf{r})$ , выполняется соотношение

$$u^* \Delta u - u \Delta u^* = 0, \text{ или } \operatorname{div} (u^* \nabla u - u \nabla u^*) = 0, \quad (26.1)$$

которое вытекает непосредственно из (24.1). Из (26.1) следует, что величину

$$\mathcal{S} = \frac{a}{2ik} (u^* \nabla u - u \nabla u^*) = \frac{a}{k} \operatorname{Im} (u^* \nabla u), \quad (26.2)$$



удовлетворяющую в прозрачной среде закону сохранения

$$\operatorname{div} \mathcal{P} = 0, \quad (26.3)$$

можно при подходящем выборе нормировочного множителя  $a$  интерпретировать как плотность потока энергии<sup>1)</sup>.

Найдем среднюю плотность потока энергии однократно рассеянного поля

$$\bar{\mathcal{P}}_1 = \frac{a}{2ik} \langle u_1^* \nabla u_1 - u_1 \nabla u_1^* \rangle. \quad (26.4)$$

Дифференцируя выражение (25.7), получаем

$$\nabla u_1(\mathbf{r}) = \frac{k_0^3}{4\pi} \int_{-\infty}^{\infty} M(\mathbf{r}') \tilde{\varepsilon}(\mathbf{r}') u_0(\mathbf{r}') \left( ik - \frac{1}{|\mathbf{r} - \mathbf{r}'|} \right) \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|} \frac{e^{ik|\mathbf{r} - \mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} d^3 r'. \quad (26.5)$$

Единичный вектор  $\mathbf{n}' \equiv (\mathbf{r} - \mathbf{r}')/|\mathbf{r} - \mathbf{r}'|$ , направленный из точки рассеяния  $\mathbf{r}'$  в точку наблюдения  $\mathbf{r}$ , можно при помощи разложения (25.11) преобразовать к виду

$$\mathbf{n}' = \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|} \approx \frac{\mathbf{r} - \mathbf{r}'}{r - (\mathbf{n} \cdot \mathbf{r}')} \approx \mathbf{n}_{s0} - \frac{\mathbf{r}'_{\perp}}{r}, \quad \mathbf{r}'_{\perp} = \mathbf{r}' - \mathbf{n}_{s0}(\mathbf{n}_{s0} \cdot \mathbf{r}'), \quad (26.6)$$

где по-прежнему  $\mathbf{n}_{s0} = \mathbf{r}/r$ . Принимая, что точка наблюдения  $\mathbf{r}$  достаточно далека от рассеивающей области ( $r \gg L$ ), имеем  $|\mathbf{r}'_{\perp}| \ll r$  и  $\mathbf{n}' \approx \mathbf{n}_{s0}$ . Кроме того, при  $r \gg L$  заведомо выполнено неравенство  $k \gg |\mathbf{r} - \mathbf{r}'|^{-1}$ . В результате выражение (26.5) сильно упрощается:

$$\nabla u_1(\mathbf{r}) \approx ik n_{s0} u_1(\mathbf{r}),$$

а формула (26.4) дает

$$\bar{\mathcal{P}}_1 \approx a n_{s0} \langle |u_1|^2 \rangle = a n_{s0} \bar{I}_1. \quad (26.7)$$

Зная  $\bar{\mathcal{P}}_1$ , нетрудно найти *эффективный поперечник* (сечение) *рассеяния*  $\sigma$  единичного объема в единичный телесный угол в направлении  $\mathbf{n}_s$ <sup>2)</sup>.

По определению

$$\sigma(\mathbf{n}_s) = \sigma(\theta, \varphi) = \frac{1}{V |\mathcal{P}_0|} \frac{d\bar{\mathcal{P}}_1}{d\Omega}, \quad (26.8)$$

<sup>1)</sup> Для звуковых волн, если  $u$  — потенциал скорости, множитель  $a$  равен  $\rho \omega^2 / 2c$  ( $\rho$  — плотность жидкости,  $c$  — скорость звука), а для электромагнитных полей  $a = c/8\pi$ , где  $c$  — скорость электромагнитной волны, если под  $u$  понимать напряженность электрического поля.

<sup>2)</sup> Для эффективного поперечника рассеяния часто используется сокращение ЭПР.

где  $d\bar{P}_1$  — средняя мощность, рассеиваемая в телесный угол  $do$  в направлении  $\mathbf{n}_s$ ,  $V$  — рассеивающий объем, а  $|\mathcal{S}_0|$  — модуль плотности потока энергии в первичной волне. Через  $\theta$  и  $\varphi$  обозначены полярный и азимутальный углы, отвечающие направлению  $\mathbf{n}_s$  (рис. 28).

Мощность  $d\bar{P}_1$  равна

$$d\bar{P}_1 = |\bar{\mathcal{S}}_1| r^2 do = \bar{a} I_1 r^2 do, \quad (26.9)$$

или, поскольку при выполнении условия (25.38) средняя интенсивность  $\bar{I}_1$  дается выражением (25.34),

$$d\bar{P}_1 = \frac{\pi a k_0^4 I_0(0) V}{2} \Phi_s(\mathbf{q}) do. \quad (26.10)$$

Плотность же потока энергии в первичной волне, будь то плоская, сферическая или направленная сферическая волна, согласно (26.2), равна (по модулю)  $|\mathcal{S}_0| = a I_0(0)$ , так что (26.8) дает

$$\sigma(\theta, \varphi) = 1/2 \pi k_0^4 \Phi_s(\mathbf{q}). \quad (26.11)$$

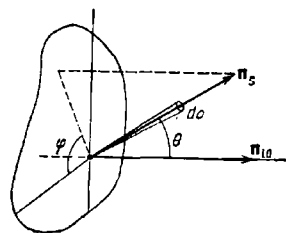


Рис. 28.

В соответствии со свойствами преобразования Фурье пространственный спектр (25.23) практически не зависит от  $\mathbf{q}$ , если корреляционная функция  $\psi_s(\rho)$  отличается от нуля только в малой области  $\rho \leq l_e \ll \lambda$ . В этом случае мелкомасштабных неоднородностей рассеяние изотропно:

$$\sigma(\theta, \varphi) \approx \text{const} = 1/2 \pi k_0^4 \Phi_s(0). \quad (26.12)$$

Напротив, в случае крупномасштабных неоднородностей ( $l_e \gg \lambda$ ) спектральная плотность  $\Phi_s(\mathbf{q})$  быстро уменьшается с ростом  $q$ , т. е. с ростом угла рассеяния  $\theta$ , что отвечает преимущественному рассеянию вперед. Сектор углов  $\theta$ , в котором сосредоточено излучение, можно оценить из условия  $q l_e \leq 1$  или  $\theta \leq 1/k l_e$ . Ниже мы проиллюстрируем эти особенности рассеяния несколькими примерами.

Величину (26.8) называют также дифференциальным сечением рассеяния — в отличие от полного поперечника рассеяния

$$\sigma_0 = \oint \sigma(\theta, \varphi) do = 1/2 \pi k_0^4 \oint \Phi_s(\mathbf{q}) do, \quad (26.13)$$

который представляет собой отношение средней полной рассеянной мощности  $\bar{P}_1 = r^2 \oint \bar{\mathcal{S}}_1 do$  к плотности потока энергии в

первичной волне в расчете на единицу объема:

$$\sigma_0 = \frac{\bar{P}_1}{V|\mathcal{S}_0|} = \frac{r^2 \oint |\bar{\mathcal{S}}_1| d\Omega}{V|\mathcal{S}_0|}. \quad (26.14)$$

Как  $\sigma$ , так и  $\sigma_0$  имеют размерность обратной длины, сечение же рассеяния всего объема  $V$ , равное  $\sigma V$ , измеряется в единицах площади.

Согласно (26.8) и (26.9) для средней плотности потока рассеянной энергии  $\bar{\mathcal{S}}_1$  и для средней интенсивности  $\bar{I}_1$  имеем

$$\bar{\mathcal{S}}_1 = \frac{n_{20}}{r^2} |\mathcal{S}_0| V \sigma(\theta, \varphi), \quad \bar{I}_1 = \frac{1}{r^2} I_0 V \sigma(\theta, \varphi). \quad (26.15)$$

В оптике  $\sigma(\theta, \varphi)$  обычно называют *индикатрисой рассеяния*, причем чаще всего имеется в виду величина  $f(\theta, \varphi)$  — значение  $\sigma(\theta, \varphi)$ , отнесенное либо к максимальному значению поперечника  $\sigma_{\max}$ , либо к полному сечению  $\sigma_0$ :

$$f(\theta, \varphi) = \frac{\sigma(\theta, \varphi)}{\sigma_{\max}} \quad \text{или} \quad f(\theta, \varphi) = \frac{\sigma(\theta, \varphi)}{\sigma_0}. \quad (26.16)$$

Полный поперечник рассеяния  $\sigma_0$  непосредственно связан с ослаблением первичной волны *за счет рассеяния* (так называемой *экстинкции*). Представим себе малый цилиндр с осью вдоль волнового вектора первичной волны  $\mathbf{k}_i = k\mathbf{n}_i$ . Объем цилиндра равен  $dV = dz d\Sigma$ , где  $d\Sigma$  — площадь поперечного сечения цилиндра, а  $dz$  — его длина. Мощность, приносимая первичной волной на передний торец цилиндра, равна  $P = |\mathcal{S}_0| d\Sigma$ , а на заднем торце она изменена до значения  $P + dP$ , где  $dP = -d\bar{P}_1$ , а  $d\bar{P}_1$  — мощность, рассеянная объемом  $dV$ , равная, согласно (26.14),

$$d\bar{P}_1 = \sigma_0 |\mathcal{S}_0| dV = \sigma_0 |\mathcal{S}_0| d\Sigma dz = P \sigma_0 dz. \quad (26.17)$$

Следовательно,

$$dP = -d\bar{P}_1 = -P \sigma_0 dz,$$

т. е. мощность первичной волны уменьшается по экспоненциальному закону  $P \sim e^{-\sigma_0 z}$ . По такому же закону уменьшается и плотность потока энергии:

$$|\mathcal{S}_0| \sim e^{-\sigma_0 z}. \quad (26.18)$$

Ослабление в  $e$  раз происходит на пути  $d_0 = 1/\sigma_0$ , который называется *длиной экстинкции* ( $\sigma_0$  называют также *коэффициентом экстинкции*). Произведение  $\tau = \sigma_0 z$ , где  $z$  — путь, пройденный волной в рассеивающей среде, называют *оптической толщиной* ( $\tau = 1$  на пути  $z = d_0$ ).

Теория однократного рассеяния строится в предположении, что амплитуда первичной волны  $u_0$ , а, стало быть, и плотность потока энергии  $\mathcal{S}_0$  практически постоянны в пределах рассеивающего объема. Согласно (26.18) это справедливо, если оптическая толщина, отвечающая размерам рассеивающего объема  $L \sim V^{1/3}$ , мала по сравнению с единицей:

$$\sigma_0 L \ll 1. \quad (26.19)$$

Условие (26.19) можно записать и в форме

$$\sigma_0 V \approx \sigma_0 L^3 \ll L^3, \quad (26.20)$$

т. е. полное сечение рассеяния всего объема  $\Sigma_0 \equiv \sigma_0 V$  должно быть мало по сравнению с площадью порядка  $L^2$ , вырезаемой объемом  $V$  из фронта первичной волны.

Приведем несколько примеров вычисления поперечника рассеяния  $\sigma(\theta, \varphi)$ .

1. Неоднородности с изотропной гауссовой корреляционной функцией:

$$\psi_e(r) = \sigma_e^2 e^{-r^2/2l_e^2}. \quad (26.21)$$

Пространственная спектральная плотность флуктуаций в рассматриваемом случае равна (см. (3.12))

$$\Phi_e(q) = \frac{\sigma_e^2 l_e^3}{(2\pi)^{3/2}} \exp\left(-\frac{q^2 l_e^2}{2}\right) = \frac{\sigma_e^2 l_e^3}{(2\pi)^{3/2}} \exp[-2k^2 l_e^2 \sin^2(1/2\theta)], \quad (26.22)$$

так что для эффективного поперечника рассеяния единичного объема по формуле (26.11) находим

$$\sigma(\theta) = \frac{\sigma_e^2 k_0^4 l_e^3}{4 \sqrt{2\pi}} \exp[-2k^2 l_e^2 \sin^2(1/2\theta)]. \quad (26.23)$$

Полный поперечник рассеяния  $\sigma_0$  вычисляется интегрированием  $\sigma(\theta)$  по единичной сфере ( $do = \sin\theta d\theta d\varphi$ ) и равен

$$\sigma_0 = \frac{\sqrt{2\pi}}{4} \frac{k_0^4 l_e^3 \sigma_e^2}{\varepsilon} \left(1 - e^{-2k^2 l_e^2}\right). \quad (26.24)$$

Условие (26.19) применимости приближения однократного рассеяния принимает в случае (26.21) вид (для простоты полагаем  $\bar{\varepsilon} = 1$ , так что  $k = k_0$ )

$$\sigma_e^2 k_0 L \ll \frac{2 \sqrt{2}}{\pi k_0 l_e [1 - \exp(-2k_0^2 l_e^2)]} \equiv \mathcal{F}(k_0 l_e). \quad (26.25)$$

Это условие ограничивает величину произведения дисперсии

флуктуаций  $\sigma_e^2$  на  $k_0 L$ . С ростом дисперсии  $\sigma_e^2$  или с увеличением  $L$  неравенство (26.25) рано или поздно нарушается. Функция

$\mathcal{F}(k_0 l_e)$  при  $k_0 l_e \rightarrow 0$  растет, как  $\sqrt{\frac{2}{\pi}} (k_0 l_e)^{-1/2}$ , а при  $k_0 l_e \rightarrow \infty$

убывает, как  $2\sqrt{2}/\pi k_0 l_e$ . Следовательно, чем больше радиус корреляции  $l_e$ , тем жестче ограничена величина  $\sigma_e^2 k_0 L$ . Поэтому рассеяние на крупных неоднородностях описывается борновским приближением лишь на сравнительно малых дистанциях. Увеличение же  $L$  требует учета многократного рассеяния. Соответствующие методы будут рассмотрены в гл. V—VII.

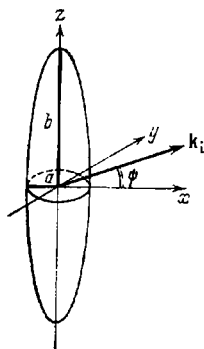


Рис. 29.

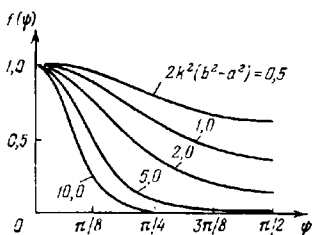


Рис. 30.

2. Обратное рассеяние на анизотропных (анизотропных) флуктуациях с гауссовой корреляционной функцией:

$$\psi_e(\mathbf{r}) = \psi_e(x, y, z) = \sigma_e^2 \exp\left(-\frac{x^2 + y^2}{2a^2} - \frac{z^2}{2b^2}\right). \quad (26.26)$$

Спектральная плотность таких флуктуаций дается выражением (3.16). Будем считать, что  $b > a$ , т. е. что большая ось эллипсоидальных неоднородностей направлена по оси  $z$ .

Пусть волновой вектор первичной волны  $\mathbf{k}_i$  лежит в плоскости  $(x, z)$  и составляет угол  $\psi$  с осью  $x$  (рис. 29), так что  $\mathbf{k}_i = (k \cos \psi, 0, k \sin \psi)$ . При обратном рассеянии  $\mathbf{q} = -2\mathbf{k}_i$  и, следовательно,

$$\begin{aligned} \sigma_{\text{обр}}(\psi) &= \frac{\pi}{2} k_0^4 \Phi_e(-2\mathbf{k}_i) = \\ &= \frac{\pi}{2} k_0^4 \frac{\sigma_e^2 a^2 b}{(2\pi)^{3/2}} \exp[-2k^2(a^2 \cos^2 \psi + b^2 \sin^2 \psi)]. \end{aligned} \quad (26.27)$$

Рассмотрим зависимость сечения обратного рассеяния  $\sigma_{обр}$  от угла  $\psi$ , который иногда называют *ракурсным* углом. При  $\psi=0$  (так называемое «ракурсное условие») волновой вектор первичной волны  $k_i$  перпендикулярен к большой оси неоднородностей. Индикатриса рассеяния выражается формулой

$$f(\psi) = \frac{\sigma_{обр}(\psi)}{\sigma_{обр}(0)} = \exp \{-2k^2(b^2 - a^2) \sin^2 \psi\}. \quad (26.28)$$

Она имеет максимум при  $\psi=0$  (рис. 30) и минимум при  $\psi = \pm \pi/2$  ( $k_i$  параллелен большой оси неоднородностей).

Этот эффект (так называемой *ракурсной чувствительности*) особенно отчетливо выражен при сильно вытянутых неоднородностях, когда  $k\sqrt{b^2 - a^2} \gg 1$ . В этом случае индикатрису (26.28) можно аппроксимировать выражением

$$f(\psi) \approx \exp \{-2k^2(b^2 - a^2) \psi^2\}, \quad (26.29)$$

т. е. узкой гауссовой кривой с шириной

$$\Delta\psi \sim 1/k\sqrt{b^2 - a^2} \ll 1. \quad (26.30)$$

**3. Рассеяние на турбулентных флуктуациях в атмосфере.** В инерционном интервале волновых чисел  $2\pi/L_0 \ll q \ll 2\pi/l_0$ , соответствующем колмогоровскому закону  $2/3$  (§ 4), спектральная плотность турбулентных флуктуаций описывается степенным законом

$$\Phi_\varepsilon(q) = \mathcal{A} C_\varepsilon^2 q^{-11/3}, \quad \mathcal{A} = 0,033. \quad (26.31)$$

Для такого спектра поперечник рассеяния единицы объема в единичный телесный угол равен (полагаем  $\bar{\varepsilon} = 1$ )

$$\sigma(\theta) = 2^{-1/2} \pi \mathcal{A} C_\varepsilon^2 k_0^{-1/3} (\sin^2 \theta)^{-11/3} = 0,016 C_\varepsilon^2 k_0^{1/3} (\sin^2 \theta)^{-11/3}. \quad (26.32)$$

Полный поперечник рассеяния оказывается в этом случае бесконечным, так как интеграл

$$\sigma_s = \oint \sigma d\omega = 2\pi \int_0^\pi \sigma(\theta) \sin \theta d\theta$$

расходится при малых углах рассеяния  $\theta$ . Объясняется это тем, что выражение для спектральной плотности (26.31), отвечающей инерционному интервалу  $2\pi/L_0 \ll q \ll 2\pi/l_0$ , непригодно при малых значениях  $q = 2k \sin^2 \theta$ . Расчет полного сечения рассеяния для другой модели спектра (4.20), принимающего при  $q \rightarrow 0$  конечные значения, приведен в задаче 2.

### § 27. Пространственная корреляция и вероятностные распределения рассеянного поля

Возвращаясь к выражению (25.24) для функции пространственной корреляции рассеянного поля, обратим внимание на сходство этой формулы с формулой (12.20), которая описывает пространственную корреляцию поля системы независимых излучателей. В обеих формулах зависимость от координат точек наблюдения  $\mathbf{r}_1$  и  $\mathbf{r}_2$  входит через один и тот же нодынтегральный множитель  $\exp(iks)/|\mathbf{r}_1 - \mathbf{R}| |\mathbf{r}_2 - \mathbf{R}|$ , где

$$s = |\mathbf{r}_1 - \mathbf{R}| + |\mathbf{r}_2 - \mathbf{R}|. \quad (27.1)$$

Это не случайное совпадение. Как уже было отмечено (§ 24), нахождение *однократно* рассеянного поля — задача, относящаяся к схеме 2), т. е. к распространению волн в случайно-неоднородной среде, фактически сводится к задаче типа 1), к возбуждению полей заданными случайными источниками. Согласно (24.11) эти источники непрерывно распределены по объему  $V$ . Однако, в силу конечности радиуса корреляции  $l_e$  неоднородностей  $\bar{\epsilon}$ , непрерывное распределение источников равносильно конечному числу (порядка  $V/l_e^3$ ) дискретных некоррелированных источников. Именно в этом и лежит причина сходства формул (25.24) и (12.20).

Опираясь на это сходство, можно сделать ряд качественных и количественных заключений о характере пространственной корреляции рассеянного поля. Так, можно утверждать, что внутри рассеивающего объема, а также вблизи него (т. е. при  $r \leq L$ ) радиус корреляции поля  $l_n$  порядка длины волны  $\lambda$ , если неоднородности мелкие ( $l_e \ll \lambda$ ), и порядка  $l_e$ , если  $l_e \gg \lambda$ . Оба утверждения вытекают из оценки

$$l_n \sim \lambda/\gamma, \quad (27.2)$$

где  $\gamma$  — видимый угловой размер области, занятой источниками. В самом деле, в случае мелкомасштабных неоднородностей, которые рассеивают изотропно, угол  $\gamma$  сравним с  $\pi$ , так что  $l_n \sim \lambda/\pi \sim \lambda$ . Крупномасштабные же неоднородности имеют узкую индикатрису ( $\gamma \sim \lambda/l_e$ ), ширина которой определяет эффективный угловой размер области, занятой источниками. В результате здесь  $l_n \sim \lambda/\gamma \sim l_e$ .

По мере удаления от рассеивающего объема происходит некоторое упорядочение поля, что выражается в увеличении масштабов пространственной корреляции. При  $r \gg L$  угловой размер рассеивающего объема  $\gamma$  становится величиной порядка  $L/r \ll 1$ , и в результате поперечный (по отношению к направлению рассеяния  $\mathbf{n}_s$ ) радиус корреляции превышает длину волны в  $r/L$  раз:

$$l_{\perp} \sim \lambda r/L. \quad (27.3)$$

При приближении к началу фраунгоферовой зоны,  $r \sim kL^2$ , поперечный радиус корреляции увеличивается до диаметра рассеивающего объема,  $l_{\perp} \sim L$ , а при  $r \gg kL^2$  превышает  $L$ .

Что касается продольного радиуса корреляции  $l_{\parallel}$ , то его можно оценить по формулам задачи 13 к гл. II: в пределах ближней зоны рассеивающего объема ( $r \ll kL^2$ )

$$l_{\parallel} \sim \frac{\lambda}{r^3} \sim \frac{\lambda r^2}{L^3} \quad (r \ll kL^2), \quad (27.4)$$

и, начиная с расстояний  $r \sim kL^2$ , корреляция простирается до бесконечности:

$$l_{\parallel} = \infty \quad (r \gg kL^2). \quad (27.5)$$

Сказанное можно частично проиллюстрировать на примере пространственной функции корреляции (25.24) при условиях, что точки наблюдения находятся в зоне Фраунгофера,  $r \gg kL^2$ , а рассеивающий объем заполнен статистически однородными флуктуациями. При этих предположениях можно заменить  $ks$  приближенным значением

$$ks \approx kr_1 - kr_2 + k \Delta n_s R, \quad (27.6)$$

где  $\Delta n_s = n_{s1} - n_{s2} = r_1/r_1 - r_2/r_2$ , и вынести за знак интеграла в (25.9) все множители, кроме  $M(R)$  и  $e^{ik \Delta n_s R}$ :

$$\psi_s(r_1, r_2) = \frac{\pi k_0^4 |A|^3}{2} \frac{\Phi_s(q_0)}{r^2 r_0^2} e^{ik(r_1 - r_2)} \int_{-\infty}^{\infty} M(R) e^{ik \Delta n_s R} d^3 R. \quad (27.7)$$

Мы заменили здесь произведение  $|\mathbf{r}_1 - \mathbf{R}| |\mathbf{r}_2 - \mathbf{R}|$  на  $r^2 = (\mathbf{r}_1 + \mathbf{r}_2)^2/4$ , сделав дополнительное предположение о малости расстояния между точками наблюдения  $|\mathbf{r}_1 - \mathbf{r}_2|$  по сравнению с расстоянием  $r$  от центра рассеивающего объема до «центра тяжести» точек наблюдения.

Интеграл в (27.7) представляет собой дельтаобразную функцию. Если ввести обозначение

$$\delta_V(\mathbf{x}) = \frac{1}{8\pi^3} \int_{-\infty}^{+\infty} M(R) \exp(i\mathbf{x} \cdot \mathbf{R}) d^3 R, \quad (27.8)$$

то формула (27.7) примет вид

$$\psi_s(r_1, r_2) = \frac{\pi k_0^4 |A|^3 \Phi_s(q_0) e^{ik(r_1 - r_2)}}{2r^2 r_0^2} 2\pi^3 \delta_V(k \Delta n_s). \quad (27.9)$$

Функция  $\delta_V(\mathbf{x})$  равна  $V/8\pi^3$  при  $\mathbf{x} = 0$  и, в соответствии со свойствами преобразования Фурье, спадает до малых значений при  $\mathbf{x} \geq 2\pi/L$ , где  $L$  — поперечник объема  $V$ . При  $V \rightarrow \infty$  она переходит в  $\delta(\mathbf{x})$ .



Разделив корреляционную функцию (27.9) на среднюю интенсивность (25.34), получаем коэффициент корреляции

$$K_n(r_1, r_2) \equiv \frac{\Psi_n(r_1, r_2)}{\sqrt{I_1(r_1) I_1(r_2)}} = \frac{8\pi^2 e^{ik(r_1-r_2)}}{V} \delta_V(k \Delta n_s). \quad (27.10)$$

Поскольку  $\delta_V(k \Delta n_s)$  заметно отличается от нуля только при  $k|\Delta n_s| \leq 2\pi/L$ , значения рассеянного поля в точках  $r_1$  и  $r_2$  становятся некоррелированными при  $\Delta n_s \geq 2\pi/kL = \lambda/L$ .

Величина  $|\Delta n_s|$  приблизительно равна углу  $\Delta\theta$  между единичными векторами  $\mathbf{n}_{s1}$  и  $\mathbf{n}_{s2}$  (рис. 23), так что полученная оценка определяет «угол корреляции»

$$\Delta\theta_k \sim \frac{\lambda}{L}. \quad (27.11)$$

Но вместе с тем  $\Delta\theta_k \sim l_\perp/r$ , так что для поперечного радиуса корреляции из (27.11) следует прежняя оценка (27.3).

Что касается продольной корреляции, то при расположении точек наблюдения на одной прямой, когда  $\Delta n_s = 0$ , коэффициент корреляции дается выражением

$$K_n(r_1, r_2) = e^{ik(r_1-r_2)},$$

т. е. равен по модулю единице при любых  $r_1$  и  $r_2$ , лежащих в дальней зоне. Это означает, в согласии с (27.5), что для амплитуды рассеянного поля  $A_1 \equiv |u_1|$  продольный радиус корреляции бесконечен,  $l_\parallel = \infty$ , а разность фаз полей в точках  $r_1$  и  $r_2$ , лежащих на одной прямой, равна разности оптических путей  $k(r_1 - r_2)$ .

Обратим внимание на то, что в формулу (27.10) входят только геометрические параметры задачи  $r$  и  $L$ , а также длина волны первичного поля  $\lambda$ , но не входят статистические характеристики флуктуаций  $\tilde{\epsilon}$ . Это говорит о том, что корреляционные характеристики рассеянного поля вытекают из чисто динамических соображений и лишь косвенно связаны со статистическим описанием.

Действительно, синусоидальная дифракционная решетка конечной длины  $L$  и с периодом  $\Lambda_d$  даст в направлении  $\theta$  (см. (25.33)) волновой пучок конечной угловой ширины  $\Delta\theta \sim \lambda/L$ . Поскольку в спектре флуктуаций представлены различные пространственные гармоники, рассеяние на данный угол обусловлено не только той объемной решеткой, которая точно удовлетворяет условию Вульфа — Брегга (25.33), но и близкими по размерам и ориентациям решетками, для которых рассматриваемое направление лежит в пределах главных дифракционных максимумов ширины  $\Delta\theta \sim \lambda/L$ . Таким образом, пока угловое расстояние между точками наблюдения меньше  $\Delta\theta \sim \lambda/L$ , рассеяние обусловлено

вполне определенной группой дифракционных решеток. Напротив, при  $\Delta\theta \geq \lambda/L$  поля, рассеянные в направлениях  $\mathbf{n}_{s1} = \mathbf{r}_1/r_1$  и  $\mathbf{n}_{s2} = \mathbf{r}_2/r_2$ , обусловлены уже *различными* группами решеток (с неперекрывающимися главными максимумами), что и приводит к исчезновению пространственной корреляции.

Обратимся теперь к вероятностным распределениям рассеянного поля. В приближении однократного рассеяния поле  $u_1(\mathbf{r})$  выражается интегралом (24.11) от произведения некоторой детерминированной функции на случайную функцию  $\tilde{\epsilon}(\mathbf{r}')$ . Поскольку линейные размеры  $L$  рассеивающего объема по предположению велики по сравнению с радиусом корреляции  $l_e$  флуктуаций  $\tilde{\epsilon}$ , можно утверждать, что в силу центральной предельной теоремы теории вероятностей закон распределения для рассеянного поля близок к нормальному.

Найдем параметры, характеризующие совместные функции распределения вещественной и мнимой частей комплексного поля:

$$u_1(\mathbf{r}) = U(\mathbf{r}) + iV(\mathbf{r}). \quad (27.12)$$

Так как величины  $U_j \equiv U(\mathbf{r}_j)$  и  $V_k \equiv V(\mathbf{r}_k)$  распределены по нормальному закону, достаточно получить матрицу вторых моментов этих величин, которая совпадает с корреляционной матрицей, поскольку среднее значение рассеянного поля, а значит, и средние значения  $\langle U_j \rangle$  и  $\langle V_k \rangle$  равны нулю. Билинейные средние  $\langle U_j V_k \rangle$  можно выразить тогда через корреляционные функции комплексного поля

$$\begin{aligned} \Psi_u(\mathbf{r}_j, \mathbf{r}_k) &= \langle u_1(\mathbf{r}_j) u_1^*(\mathbf{r}_k) \rangle, \\ \tilde{\Psi}_u(\mathbf{r}_j, \mathbf{r}_k) &= \langle u_1(\mathbf{r}_j) u_1(\mathbf{r}_k) \rangle. \end{aligned} \quad (27.13)$$

Первая из этих функций была вычислена в § 25 и рассмотрена выше. Покажем, что вторая корреляционная функция  $\tilde{\Psi}_u$  почти всюду мала по сравнению с  $\Psi_u$ . С этой целью оценим величину  $\langle u_1^2(\mathbf{r}) \rangle = \tilde{\Psi}_u(\mathbf{r}, \mathbf{r})$  в зоне Фраунгофера. Согласно (25.39),

$$\begin{aligned} \langle u_1^2(\mathbf{r}) \rangle &= \frac{k_0^4 A^2 e^{2ikr}}{(4\pi r r_0)^2} \int_{-\infty}^{\infty} M(\mathbf{r}') d^3 r' \int_{-\infty}^{\infty} M(\mathbf{r}'') d^3 r'' \Psi_e(\mathbf{r}' - \mathbf{r}'') e^{-iq(\mathbf{r}' + \mathbf{r}'')} = \\ &= -\frac{k_0^4 A^2 e^{2ikr}}{(4\pi r r_0)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d^3 \rho d^3 R M(\mathbf{R} + \boldsymbol{\rho}/2) M(\mathbf{R} - \boldsymbol{\rho}/2) \Psi_e(\boldsymbol{\rho}) e^{-2iq\mathbf{R}}. \end{aligned} \quad (27.14)$$

Заменяя  $M(\mathbf{R} + \boldsymbol{\rho}/2) M(\mathbf{R} - \boldsymbol{\rho}/2)$  на  $M(\mathbf{R})$ , получаем

$$\langle u_1^2(\mathbf{r}) \rangle = \frac{k_0^4 A^2 e^{2ikr}}{(4\pi r r_0)^2} (2\pi)^3 \Phi_e(0) \delta_V(2\mathbf{q}), \quad (27.15)$$

где функция  $\delta_V(\mathbf{x})$  определена выражением (27.8).

При рассеянии вперед ( $q=0$ ) средний квадрат поля  $\langle u_1^2 \rangle$  с точностью до фазы совпадает со средней интенсивностью  $\bar{I}_1 = \langle |u_1|^2 \rangle$ , поскольку  $\delta_V(0) = V/8\pi^3$ . Однако при  $|2q| \geq 2\pi/L$  значения функции  $\delta_V(2q)$  резко уменьшаются по сравнению с  $\delta_V(0)$ , т. е.

$$|\langle u_1^2(\mathbf{r}) \rangle| \ll \bar{I}_1 - \langle |u_1^2(\mathbf{r})| \rangle, \quad q \geq \pi/L. \quad (27.16)$$

Так как  $q = 2k \sin \frac{1}{2}\theta$ , неравенство  $q \geq \pi/L$  выполняется вне узкого конуса с раствором  $\theta \sim \lambda/2L$ , откуда и следует, что при рассеянии на не слишком малые углы ( $\theta \geq \lambda/2L$ )

$$|\tilde{\Psi}_u| \ll |\Psi_u|. \quad (27.17)$$

Практически это означает, что вне указанного конуса вторую корреляционную функцию  $\tilde{\Psi}_u$  можно положить равной нулю:

$$\tilde{\Psi}_u(\mathbf{r}_j, \mathbf{r}_k) \approx 0. \quad (27.18)$$

Этот результат получен для дальней зоны рассеивающего объема ( $r \gg kL^2$ ), но в ближней зоне ( $r \ll kL^2$ ) он справедлив и по-прежнему. Действительно, если разбить рассеивающий объем на отдельные элементы с линейными размерами, малыми по сравнению с  $L$ , но большими по сравнению с радиусом корреляции неоднородностей, и поместить точку наблюдения в зону Фраунгофера каждого из таких элементов, то результирующий средний квадрат поля  $\langle u_1^2 \rangle$  можно получить, суммируя выражения типа (27.15), поскольку флуктуации  $\epsilon$  в различных элементах статистически независимы. Ясно, что суммирование величин, содержащих осциллирующий множитель  $e^{2ikr}$ , может привести только к уменьшению  $\langle u_1^2 \rangle$  по сравнению с  $\bar{I}_1 = \langle |u_1|^2 \rangle$  даже внутри конуса с углом раствора  $\theta \sim \lambda/2L$ .

Воспользовавшись этим, применим для вычисления моментов  $\langle U_j V_k \rangle$  формулы (2.14), которые при равенстве нулю второй корреляционной функции дают

$$\begin{aligned} \langle U_1 U_2 \rangle &= \langle V_1 V_2 \rangle = \frac{1}{2} \operatorname{Re} \Psi_u(\mathbf{r}_1, \mathbf{r}_2), \\ \langle V_1 U_2 \rangle &= -\langle U_1 V_2 \rangle = \frac{1}{2} \operatorname{Im} \Psi_u(\mathbf{r}_1, \mathbf{r}_2). \end{aligned} \quad (27.19)$$

Положив в (27.19)  $\mathbf{r}_1 = \mathbf{r}_2 = \mathbf{r}$ , находим, в частности, что

$$\langle U^2 \rangle = \langle V^2 \rangle = \frac{1}{2} \bar{I}_1, \quad \langle U(\mathbf{r}) V(\mathbf{r}) \rangle = 0. \quad (27.20)$$

Последующие вычисления статистических моментов амплитуды и фазы производятся так же, как в ч. 1, §§ 25 и 44.

## § 28. Рассеяние на нестационарных неоднородностях

1. Временная функция корреляции. Поле, рассеянное на неоднородностях  $\tilde{\epsilon}(t, \mathbf{r})$ , зависящих от времени, создает на выходе антенны отклик, аналогичный (25.48):

$$v(t, \mathbf{r}) = e^{-i\omega t} \int \mathcal{P}(\mathbf{r}') e^{ik(|\mathbf{r}-\mathbf{r}'|+|\mathbf{r}'-\mathbf{r}_0|)} \tilde{\epsilon}\left(t - \frac{|\mathbf{r}-\mathbf{r}'|}{c}, \mathbf{r}'\right) d^3r'. \quad (28.1)$$

В отличие от (25.48), мы восстановили здесь множитель  $e^{-i\omega t}$ , который для краткости ранее опускали. Волновое возмущение, возникающее при рассеянии в точке  $\mathbf{r}'$ , достигает точки наблюдения  $\mathbf{r}$  за конечное время  $\Delta t = |\mathbf{r}-\mathbf{r}'|/c$  (для простоты считаем, что  $\tilde{\epsilon} = 1$ ). Поэтому под интеграл входят значения  $\tilde{\epsilon}$  не в момент  $t$ , а в предшествующий момент  $t' = t - \Delta t$ . Функция  $\mathcal{P}(\mathbf{r}')$  по-прежнему дается выражением (25.47).

Вычислим временную корреляционную функцию рассеянного поля в предположении, что флуктуации  $\tilde{\epsilon}$  статистически *однородны* и *стационарны*, так что пространственно-временная корреляционная функция неоднородностей зависит лишь от разностей (пространственных и временных) переменных:

$$\psi_{\epsilon}(t', \mathbf{r}'; t'', \mathbf{r}'') = \langle \tilde{\epsilon}(t', \mathbf{r}') \tilde{\epsilon}(t'', \mathbf{r}'') \rangle = \psi_{\epsilon}(t' - t'', \mathbf{r}' - \mathbf{r}''). \quad (28.2)$$

Из (28.1) сразу же видно, что временная корреляционная функция отклика на выходе антенны

$$\psi_v(t', t'') = \langle v(t') v^*(t'') \rangle$$

тоже зависит только от разности  $\tau = t' - t''$ :

$$\psi_v(\tau) = e^{-i\omega\tau} \int_{-\infty}^{\infty} \mathcal{P}(\mathbf{r}') \mathcal{P}^*(\mathbf{r}'') e^{i\Psi(\mathbf{r}', \mathbf{r}'')} \psi_{\epsilon}(\tau - \Delta\tau, \mathbf{r}' - \mathbf{r}'') d^3r' d^3r'', \quad (28.3)$$

где

$$\Psi(\mathbf{r}', \mathbf{r}'') = k \{ |\mathbf{r} - \mathbf{r}'| + |\mathbf{r}' - \mathbf{r}_0| - |\mathbf{r} - \mathbf{r}''| - |\mathbf{r}'' - \mathbf{r}_0| \}.$$

Таким образом, в случае стационарных флуктуаций  $\tilde{\epsilon}$  отклик  $v(t)$ , как и само рассеянное поле  $u_1(t)$ , является стационарным случайным процессом.

Входящая в (28.3) величина

$$\Delta\tau = (|\mathbf{r}' - \mathbf{r}| - |\mathbf{r}'' - \mathbf{r}|)/c$$

представляет собой разность временных задержек возмущений, пришедших из точек  $\mathbf{r}'$  и  $\mathbf{r}''$ . В силу того, что пространственная корреляция неоднородностей простирается на расстояния  $|\mathbf{r}' - \mathbf{r}''| \leq l_{\epsilon}$ , разность задержек  $\Delta\tau$  фактически не превышает времени  $l_{\epsilon}/c$ , за которое волна проходит одну неоднородность

( $\Delta\tau \leq l_e/c$ ). Интервал  $\Delta\tau$  обычно значительно меньше  $\tau_e$  — времени корреляции флуктуаций  $\tilde{\epsilon}$ :

$$\Delta\tau \sim l_e/c \ll \tau_e. \quad (28.4)$$

Поэтому без большой ошибки можно при этом условии положить в (28.3)  $\Delta\tau = 0$ , и тогда

$$\psi_e(\tau) = e^{-i\omega\tau} \int_{-\infty}^{\infty} \mathcal{P}(\mathbf{r}') \mathcal{P}^*(\mathbf{r}'') e^{i\mathbf{W}(\mathbf{r}', \mathbf{r}'')} \psi_e(\tau, \mathbf{r}' - \mathbf{r}'') d^3r' d^3r''. \quad (28.5)$$

Выражение (28.5) отвечает так называемому *квазистационарному* приближению, при котором считается, что рассеяние на отдельных неоднородностях происходит так, как если бы неоднородности покоились, а зависимость  $\psi_e$  от времени восстанавливается уже в окончательной формуле (28.5).

Дальнейший анализ временных флуктуаций целесообразно провести отдельно для следующих двух случаев: а) в среднем покоящаяся случайно-неоднородная среда и б) случайно-неоднородная среда, перемещающаяся в среднем равномерно со скоростью  $\mathbf{v}$ , которую принято называть *скоростью дрейфа*.

2. Рассеяние в отсутствие регулярного дрейфа. Для среды, в среднем неподвижной, можно почти без изменений повторить выкладки § 25, т. е. перейти к переменным интегрирования  $\mathbf{R} = (\mathbf{r}_1 + \mathbf{r}_2)/2$ ,  $\boldsymbol{\rho} = \mathbf{r}_1 - \mathbf{r}_2$  и разложить показатель экспоненты  $\Psi(\mathbf{r}', \mathbf{r}'')$  в ряд по степеням  $\boldsymbol{\rho}$ :

$$\begin{aligned} \psi_e(\tau) &= e^{-i\omega\tau} \int_{-\infty}^{\infty} d^3R |\mathcal{P}(\mathbf{R})|^2 \int_{-\infty}^{\infty} d^3\rho \psi_e(\tau, \boldsymbol{\rho}) \exp(i\mathbf{q}\boldsymbol{\rho}) = \\ &= 8\pi^3 e^{-i\omega\tau} \int_{-\infty}^{\infty} d^3R |\mathcal{P}(\mathbf{R})|^2 \Phi_e(\tau, \mathbf{q}), \end{aligned} \quad (28.6)$$

где

$$\Phi_e(\tau, \mathbf{q}) = \frac{1}{8\pi^3} \int_{-\infty}^{\infty} \psi_e(\tau, \boldsymbol{\rho}) \exp(-i\mathbf{q}\boldsymbol{\rho}) d^3\rho \quad (28.7)$$

— пространственное преобразование Фурье корреляционной функции  $\psi_e(\tau, \boldsymbol{\rho})$ , а  $\mathbf{q} = \mathbf{q}(\mathbf{R})$  — вектор рассеяния, отвечающий точке  $\mathbf{R}$ .

Из выражения (28.6) следует, что время корреляции отклика на выходе антенны того же порядка, что и у флуктуаций  $\tilde{\epsilon}$ . Тот же результат можно выразить и в спектральной форме,

написав спектральную плотность стационарного процесса  $v(t)$ :

$$g_v(\Omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi_a(\tau) e^{i\Omega\tau} d\tau = \\ = \frac{1}{2\pi} \int_{-\infty}^{\infty} d^3R |\mathcal{P}(\mathbf{R})|^2 \int_{-\infty}^{\infty} d^3\rho \psi^*(\tau, \rho) \exp[i(\Omega - \omega)\tau - i\mathbf{q}\rho]. \quad (28.8)$$

Учитывая, что пространственно-временной спектр флуктуаций  $\bar{\epsilon}$  выражается формулой

$$G_e(\Omega, \mathbf{k}) = \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} d\tau \int_{-\infty}^{\infty} d^3\rho \psi_e(\tau, \rho) \exp(i\Omega\tau - i\mathbf{k}\rho),$$

можно представить частотный спектр сигнала  $v(t)$  в виде

$$g_v(\Omega) = 8\pi^3 \int d^3R |\mathcal{P}(\mathbf{R})|^2 G_e[\Omega - \omega, \mathbf{q}(\mathbf{R})]. \quad (28.9)$$

Спектр флуктуаций  $G_e(\Omega, \mathbf{k})$  обычно сосредоточен в окрестности нулевой частоты  $\Omega \approx 0$ , спектр же рассеянного сигнала  $g_v(\Omega)$  расположит вблизи частоты первичной волны  $\omega$ . Как следует из (28.9), частотный спектр рассеянного поля получается суммированием пространственно-временного спектра неоднородностей  $G_e[\Omega - \omega, \mathbf{q}(\mathbf{R})]$  по рассеивающей области с весовой функцией

$$8\pi^3 |\mathcal{P}(\mathbf{R})|^2 = \frac{\pi k_0^4 |A|^2}{2} \frac{|f_0(n_l) f_1(n_s)|^2}{|r - \mathbf{r}_l| |\mathbf{R} - \mathbf{r}_0|} M(\mathbf{R}).$$

Если условия задачи позволяют пренебречь изменением вектора рассеяния  $\mathbf{q}$  в пределах области интегрирования (согласно (25.38) это возможно при  $r \gg r_{\min} = \min(L, LLe/\lambda)$ ), то  $G_e[\Omega - \omega, \mathbf{q}(\mathbf{R})]$  можно вынести за знак интеграла со значением при  $\mathbf{q} = \mathbf{q}_0$ , где  $\mathbf{q}_0 = k(\mathbf{n}_{s0} - \mathbf{n}_{l0})$  — вектор рассеяния, отвечающий центру рассеивающей области. В результате частотный спектр поля оказывается пропорциональным  $G_e(\Omega - \omega, \mathbf{q}_0)$ :

$$g_v(\Omega) = 8\pi^3 G_e(\Omega - \omega, \mathbf{q}_0) \int_{-\infty}^{\infty} d^3R |\mathcal{P}(\mathbf{R})|^2 = \\ = G_e(\Omega - \omega, \mathbf{q}_0) \frac{\pi}{2} \frac{k_0^4 |A|^2}{r_0^2 r^2} V_{\text{эфф}}, \quad (28.10)$$

где  $V_{\text{эфф}}$  — эффективный рассеивающий объем. Таким образом, при  $r \gg r_{\min}$  можно находить форму спектра флуктуаций  $\epsilon$  непосредственно по измеренным значениям  $g_v(\Omega)$ . В случае же протяженных рассеивающих областей, когда спектры  $g_v(\Omega)$  и  $G_e$  связаны между собой интегральным соотношением (28.9), определение формы  $G_e(\Omega)$  по измерениям  $g_v(\Omega)$  затруднительно.

Спектр комплексной огибающей  $\mathcal{U}(t)$  сигнала  $v(t) = \mathcal{U}(t)e^{-i\omega t}$  сосредоточен в окрестности нулевой частоты, поскольку  $g_{\pi}(\Omega) = g_v(\Omega + \omega)$ . В частности, при  $r \gg r_{\min}$ , когда  $g_v(\Omega)$  дается выражением (28.10), имеем

$$g_{\pi}(\Omega) = G_e(\Omega, \mathbf{q}_0) \frac{\pi}{2} \frac{k_0^4 |A|^2}{r^2 r_0^2} V_{\text{эфф}}.$$

3. Рассеяние при наличии регулярного дрейфа. Нахождение функции корреляции поля, рассеянного на дрейфующих неоднородностях, несколько сложнее, чем при  $\bar{\mathbf{v}} = 0$ . При рассмотрении этого вопроса мы следуем результатам Г. С. Горелика и его сотрудников [3—6], а также анализу, приведенному в [1].

Пусть  $\psi_{\text{др}}(\tau, \rho)$  — функция корреляции неоднородностей в сопровождающей системе координат, т. е. в системе, равномерно перемещающейся вместе со средой со скоростью  $\bar{\mathbf{v}}$ . В неподвижной системе координат, связанной с наблюдателем, функция корреляции, очевидно, равна

$$\psi_e(\tau, \rho) = \psi_{\text{др}}(\tau, \rho - \bar{\mathbf{v}}\tau), \quad \rho = \mathbf{r}' - \mathbf{r}''. \quad (28.11)$$

Подставим это выражение в формулу (28.5) и введем новые переменные интегрирования  $\mathbf{R}$  и  $\rho'$ , равные

$$\mathbf{R} = 1/2(\mathbf{r}' + \mathbf{r}''), \quad \rho' = \rho - \bar{\mathbf{v}}\tau = \mathbf{r}' - \mathbf{r}'' - \bar{\mathbf{v}}\tau. \quad (28.12)$$

В этих переменных функция корреляции сигнала  $v(t)$  будет

$$\begin{aligned} \psi_v(\tau) = e^{-i\omega\tau} \iint_{-\infty}^{\infty} \mathcal{P}\left(\mathbf{R} + \frac{\rho' + \bar{\mathbf{v}}\tau}{2}\right) \mathcal{P}^*\left(\mathbf{R} - \frac{\rho' - \bar{\mathbf{v}}\tau}{2}\right) \psi_{\text{др}}(\tau, \rho') \times \\ \times \exp[i\Psi'(\tau, \rho', \mathbf{R})] d^3\rho d^3R, \end{aligned} \quad (28.13)$$

где

$$\begin{aligned} \Psi'(\tau, \rho', \mathbf{R}) = \Psi\left(\mathbf{R} + \frac{\rho' + \bar{\mathbf{v}}\tau}{2}, \mathbf{R} - \frac{\rho' - \bar{\mathbf{v}}\tau}{2}\right) = \\ = k\left(\left|\mathbf{r} - \mathbf{R} - \frac{\rho' + \bar{\mathbf{v}}\tau}{2}\right| + \left|\mathbf{R} + \frac{\rho' + \bar{\mathbf{v}}\tau}{2} - \mathbf{r}_0\right| - \right. \\ \left. - \left|\mathbf{r} - \mathbf{R} + \frac{\rho' - \bar{\mathbf{v}}\tau}{2}\right| - \left|\mathbf{R} - \frac{\rho' - \bar{\mathbf{v}}\tau}{2} - \mathbf{r}_0\right|\right). \end{aligned}$$

Поскольку корреляционная функция  $\psi_{\text{др}}(\tau, \rho')$  спадает практически до нуля при  $\rho' \gg l_e$ , разложим  $\Psi'$  в степенной ряд по  $\rho'$  и ограничимся первыми двумя членами разложения:

$$\Psi'(\tau, \rho', \mathbf{R}) = \Phi(\tau, \mathbf{R}) - \mathbf{q}'(\tau, \mathbf{R})\rho' + \dots \quad (28.14)$$

Здесь

$$\begin{aligned}\Phi(\tau, \mathbf{R}) &= \Psi'(\tau, 0, \mathbf{R}) = k(|\mathbf{r} - \mathbf{R} - \bar{\mathbf{v}}\tau/2| + \\ &\quad - |\mathbf{R} + \bar{\mathbf{v}}\tau/2 - \mathbf{r}_0| - |\mathbf{r} - \mathbf{R} + \bar{\mathbf{v}}\tau/2| - |\mathbf{R} - \bar{\mathbf{v}}\tau/2 - \mathbf{r}_0|), \\ q'(\tau, \mathbf{R}) &= \frac{k}{2} \left( \frac{|\mathbf{r} - \mathbf{R} - \bar{\mathbf{v}}\tau/2|}{|\mathbf{r} - \mathbf{R} - \bar{\mathbf{v}}\tau/2|} - \frac{|\mathbf{R} + \bar{\mathbf{v}}\tau/2 - \mathbf{r}_0|}{|\mathbf{R} + \bar{\mathbf{v}}\tau/2 - \mathbf{r}_0|} - \right. \\ &\quad \left. - \frac{|\mathbf{r} - \mathbf{R} + \bar{\mathbf{v}}\tau/2|}{|\mathbf{r} - \mathbf{R} + \bar{\mathbf{v}}\tau/2|} + \frac{|\mathbf{R} - \bar{\mathbf{v}}\tau/2 - \mathbf{r}_0|}{|\mathbf{R} - \bar{\mathbf{v}}\tau/2 - \mathbf{r}_0|} \right).\end{aligned}$$

Если подставить в (28.13) приближение (28.14), справедливое при выполнении довольно слабых условий типа (25.18), то интегрирование по  $\rho'$  даст

$$\psi_v(\tau) = 8\pi^3 e^{-i\omega\tau} \int_{-\infty}^{\infty} d^3R \mathcal{P}_+ \mathcal{P}_-^* e^{i\Phi(\tau, \mathbf{R})} \Phi_{\text{лр}}(\tau, \mathbf{q}'), \quad (28.15)$$

где  $\mathcal{P}_{\pm} \equiv \mathcal{P}(\mathbf{R} \pm \bar{\mathbf{v}}\tau/2)$ , а функция

$$\Phi_{\text{лр}}(\tau, \mathbf{x}) = \frac{1}{8\pi^3} \int_{-\infty}^{\infty} \Psi_{\text{лр}}(\tau, \rho) \exp(i\mathbf{x}\rho) d^3\rho \quad (28.16)$$

представляет собой пространственное преобразование Фурье корреляционной функции неоднородностей в сопровождающей системе координат.

Выражение (28.15) сложнее, чем (28.5). Во-первых, под знаком интеграла теперь содержится дополнительный экспоненциальный множитель  $e^{i\Phi(\tau, \mathbf{R})}$ . Во-вторых, вектор рассеяния  $\mathbf{q}'$  и произведение  $\mathcal{P}_+ \mathcal{P}_-^*$ , заменяющее теперь  $|\mathcal{P}(\mathbf{R})|^2$ , зависят от  $\tau$ . Однако если допустить, что источник первичного поля и приемный пункт удалены от рассеивающей области на расстояния, превышающие ее поперечник ( $r \gg L$  и  $r_0 \gg L$ ), то формула (28.15) упрощается. В последующих оценках мы будем писать неравенства, относящиеся только к  $r$ . Аналогичные неравенства должны выполняться и для  $r_0$ .

При  $R \gg L$  произведение  $\mathcal{P}_+ \mathcal{P}_-^*$  стремится к нулю. Следовательно, в области  $R \leq L$ , существенной для интегрирования в (28.15), отношение  $R/r \ll 1$  в силу принятого допущения  $r \gg L$ . По той же причине в (28.15) существенны только значения  $\bar{\mathbf{v}}\tau \leq L$ . Имея это в виду, разложим показатель экспоненты  $\Phi(\tau, \mathbf{R})$  в степенной ряд по малым параметрам  $R/r \leq L/r \ll 1$  и  $\bar{\mathbf{v}}\tau/r \leq L/r \ll 1$ :

$$\Phi(\tau, \mathbf{R}) = -\mathbf{q}_0 \bar{\mathbf{v}}\tau + k\mathbf{v}\mathbf{R}\tau + kO\left[\left(\frac{R}{r}\right)^2 \bar{\mathbf{v}}\tau\right] + kO\left[\frac{(\bar{\mathbf{v}}\tau)^2}{r^3}\right]. \quad (28.17)$$



Здесь  $\mathbf{v}$  — вектор, имеющий размерность частоты и пропорциональный скорости дрейфа  $\bar{\mathbf{v}}$ :

$$\mathbf{v} = \frac{1}{r} [n_{s0} [\bar{\mathbf{v}} n_{s0}]] + \frac{1}{r_0} [n_{l0} [\bar{\mathbf{v}} n_{l0}]]. \quad (28.18)$$

Вектор  $\mathbf{q}'$  мы тоже разложим в ряд, но лишь по переменной  $\tau$ :

$$\mathbf{q}'(\tau, \mathbf{R}) = \mathbf{q}(\mathbf{R}) + kO \left[ \left( \frac{\bar{v}\tau}{r} \right)^2 \right], \quad \mathbf{q}(\mathbf{R}) = k(n_s - n_l) \quad (28.19)$$

(член, линейный по  $\tau$ , здесь отсутствует).

Ниже мы убедимся, что условие  $r \gg L$  достаточно для того, чтобы сохранить в (28.17) два первых члена, а в (28.19) — только первый член разложения. Если это сделать, то (28.15) принимает вид

$$\psi_v(\tau) = 8\pi^2 e^{-i(\omega + \mathbf{q}_s \bar{\mathbf{v}})\tau} \int \mathcal{P}_+ \mathcal{P}_-^* \exp(ik\mathbf{v}\mathbf{R}\tau) \Phi_{\text{др}}(\tau, \mathbf{q}) d^3R. \quad (28.20)$$

Разумеется, при  $\bar{\mathbf{v}} \rightarrow 0$  это выражение переходит в формулу (28.5), полученную в отсутствие дрейфа.

С ростом  $\tau$  значение интеграла в (28.20) уменьшается по трем причинам. Во-первых, уменьшается произведение  $\mathcal{P}_+ \mathcal{P}_-^*$ , которое стремится к нулю при  $\bar{v}\tau \geq L$ . Во-вторых, экспоненциальная функция  $\exp(ik\mathbf{v}\mathbf{R}\tau)$  начинает заметно осциллировать в пределах существенной для интегрирования области. В-третьих, уменьшается функция  $\Phi_{\text{др}}(\tau, \mathbf{q})$ , которая описывает временную корреляцию флуктуаций в сопровождающей системе координат и стремится к нулю при  $\tau \rightarrow \infty$ .

Из условий  $\bar{v}\tau \sim L$  и  $k\mathbf{v}\mathbf{R}\tau \leq k\bar{v}L\tau/r \sim 1$  находим два характерных интервала времени, отвечающих первым двум факторам

$$\begin{aligned} \tau_1 &\sim \frac{L}{\bar{v}}, \\ \tau_2 &\sim \frac{r}{k\bar{v}L} \sim \frac{\lambda r}{\bar{v}L}. \end{aligned} \quad (28.21)$$

Характерное же время изменения функции  $\Phi_{\text{др}}(\tau, \mathbf{q})$ , т. е. временной интервал корреляции неоднородностей в сопровождающей системе координат, мы обозначим через  $\tau_3$ . По существу, это «время жизни» неоднородностей. Меньшее из трех перечисленных значений  $\tau$  определяет, очевидно, время корреляции  $\tau_x$  сигнала:

$$\tau_x = \min \{\tau_1, \tau_2, \tau_3\}. \quad (28.22)$$

При  $\bar{v} \rightarrow 0$  (отсутствие дрейфа)  $\tau_1$  и  $\tau_2$  бесконечно возрастают, а  $\tau_x = \tau_3$ .

Применим формулу (28.20) к анализу частного случая рассеяния на *замороженных* неоднородностях, которые в сопровождающей системе координат не зависят от времени ( $\tau_s = \infty$ ).

4. Рассеяние на замороженных неоднородностях. Очевидно, в лабораторной системе координат, относительно которой не меняющиеся со временем неоднородности дрейфуют со скоростью  $\bar{v}$ , флуктуации проницаемости  $\bar{\epsilon}(t, \mathbf{r})$  обладают свойством

$$\bar{\epsilon}(t, \mathbf{r}) = \bar{\epsilon}(t - \tau, \mathbf{r} - \bar{v}\tau), \quad (28.23)$$

т. е. в точку  $\mathbf{r}$  в момент времени  $t$  приходит возмущение  $\bar{\epsilon}(t', \mathbf{r}')$ , которое в предшествующий момент  $t' = t - \tau$  находилось в точке  $\mathbf{r}' = \mathbf{r} - \bar{v}\tau$  и переместилось за время  $\tau$  со скоростью  $\bar{v}$  в точку  $\mathbf{r}$ .

Как сказано, «время жизни» неоднородностей  $\tau_s$  в сопровождающей системе координат формально бесконечно:  $\tau_s = \infty$ . Практически же замороженными можно считать неоднородности, для которых

$$\tau_s \gg \tau_1, \tau_2. \quad (28.24)$$

При выполнении этого условия функцию  $\Phi_{\text{др}}(\tau, \mathbf{q})$  в (28.20) можно положить равной  $\Phi_{\text{др}}(0, \mathbf{q}) = \Phi_s(\mathbf{q})$ , поскольку значение интеграла в (28.20) делается малым раньше, чем начнется уменьшение  $\Phi_{\text{др}}(\tau, \mathbf{q})$ . Таким образом, для замороженных неоднородностей

$$\psi_s(\tau) = 8\pi^3 e^{-i(\omega + \mathbf{q}_0 \bar{v})\tau} \int_{-\infty}^{\infty} \mathcal{P}_+ \mathcal{P}_-^* \exp(ikvR\tau) \Phi_s(\mathbf{q}) d^3R. \quad (28.25)$$

Поведение этой функции корреляции существенно зависит от соотношения между  $\tau_1$  и  $\tau_2$ . Согласно (28.21) они сравнимы друг с другом при  $r \sim kL^2$ . В ближней зоне, где  $r \ll kL^2$  и одновременно, по условию,  $r \gg L$ , выполняется неравенство  $\tau_2 \ll \tau_1$ . Следовательно, время корреляции в ближней зоне совпадает с  $\tau_2$ :

$$\tau_k = \tau_2 \sim \frac{r}{kvL} \quad (L \ll r \ll kL^2). \quad (28.26)$$

Напротив, в дальней (фраунгоферовой) зоне, где  $r \gg kL^2$ , время корреляции совпадает с  $\tau_1$ , которое значительно меньше  $\tau_2$ :

$$\tau_k = \tau_1 \sim L/v \quad (r \gg kL^2). \quad (28.27)$$

Нетрудно убедиться, что в обеих зонах отбрасывание «лишних» слагаемых в разложениях (28.17) и (28.19) вполне оправдано. Например, в ближней зоне, где  $r \ll kL^2$  и  $\tau_k = \tau_2$ , четвертый член разложения (28.17) по порядку величины равен

$$k \frac{(\bar{v}\tau)^3}{r^3} \ll \frac{k}{r^3} \left( \frac{r}{kL} \right)^3 = \frac{r}{kL^3} \frac{1}{kL},$$

т. е. меньше произведения двух малых параметров. Но он мал и в дальней зоне, где  $r \gg kL^2$  и  $\tau_2 = \tau_1$ , так что

$$k \frac{(\bar{v}\tau)^2}{r^2} \ll \frac{k}{r^2} L^2 = \frac{kL^2}{r} \frac{L}{r} \ll 1.$$

Отметим, что при нарушении условия замороженности (28.24), т. е. при  $\tau_1, \tau_2 \geq \tau_3$ , отбрасывание «лишних» членов является тем более оправданным, чем меньше  $\tau_3$  по сравнению с  $\tau_1$  и  $\tau_2$ . Иными словами, единственное условие применимости формулы (28.20) выражается неравенством  $r \gg L$ .

Рассмотрим теперь более детально поведение корреляционной функции (28.25) в ближней и дальней зонах.

Для ближней зоны выражение (28.25) можно упростить, если учесть, что  $\bar{v}\tau \ll \bar{v}\tau_2 \sim r/kL \ll L$ , а при  $\bar{v}\tau \ll L$  произведение  $\mathcal{P}_+ \mathcal{P}_-$  приближенно равно  $|\mathcal{P}(R)|^2$ . Таким образом, здесь

$$\psi_v(\tau) = e^{-i(\omega + q_0 \bar{v})\tau} \int_{-\infty}^{\infty} |\mathcal{P}(R)|^2 \exp(ikvR\tau) \Phi_e(q) d^3R. \quad (28.28)$$

Для вычисления интеграла, конечно, надо конкретизировать вид спектра  $\Phi_e(q)$  и весовой функции  $|\mathcal{P}(R)|^2$ . Однако спектральный анализ выражения (28.28) можно провести и без конкретизации этих функций.

Выполнив временное преобразование Фурье (28.28), находим спектральную плотность сигнала  $v(t)$ :

$$g_v(\Omega) = 8\pi^3 \int_{-\infty}^{\infty} |\mathcal{P}(R)|^2 \Phi_e(q) \delta(\Omega - \omega - q_0 \bar{v} - kvR) d^3R. \quad (28.29)$$

Согласно (28.29) каждому элементарному объему  $d^3R$  отвечает спектральная линия на частоте  $\Omega - \omega + q_0 \bar{v} + kvR$  и с «интенсивностью»  $8\pi^3 |\mathcal{P}(R)|^2 \Phi_e(q)$ . Суперпозиция этих линий и образует частотный спектр  $g_v(\Omega)$ . Смещение частоты линии  $\Omega = \omega + q_0 \bar{v} + kvR$  относительно частоты первичного поля  $\omega$  обусловлено эффектом Доплера.

Доплеровский сдвиг частоты

$$\Omega_d = \bar{v}q_0 + kvR$$

содержит два слагаемых, из которых первое одинаково для всех точек рассеивающей области, тогда как второе меняется от точки к точке в силу изменения вектора  $q(R)$  внутри рассеивающего объема<sup>1)</sup>.

<sup>1)</sup> Более точное выражение доплеровского сдвига:  $\Omega_d = \bar{v}q(R)$ . Выражение, приведенное выше, содержит только первые члены тейлоровского разложения по  $R$ .

Ширина частотного спектра  $\Delta\Omega$  определяется, очевидно, разностью доплеровских смещений частоты на краях рассеивающего объема  $R \sim L$ :

$$\Delta\Omega \approx \Delta\Omega_{\kappa} \sim kvL \sim k\bar{v}L/r, \quad (28.30)$$

что, как и следовало ожидать, согласуется с оценкой ширины спектра  $\Delta\Omega \sim 1/\tau_{\kappa} \sim 1/\tau_2$ . Оценке (28.26) можно дать и другую интерпретацию. Как было показано в § 27, величина  $r/kL$  представляет собой в случае неподвижных неоднородностей поперечный радиус корреляции поля  $l_{\perp}$ , т. е. пространственный масштаб интерференционной картины, образующейся в результате наложения рассеянных волн. При дрейфе замороженных неоднородностей интерференционная картина будет пробегать мимо точки наблюдения со скоростью  $\bar{v}$ , так что характерное время флуктуаций поля  $l_{\perp}/\bar{v} \sim r/k\bar{v}L$  как раз совпадает с  $\tau_{\kappa} = \tau_2$ .

Оценка (28.26) пригодна, строго говоря, только при  $r \gg L$ . Если, несмотря на это, применить ее к случаю  $r \sim L$ , когда точка наблюдения находится вблизи рассеивающей области, то мы получим, что  $\tau_{\kappa} \sim 1/k\bar{v} \sim \lambda/\bar{v}$ . Это не противоречит и оценке  $\tau_{\kappa} \sim l_{\perp}/\bar{v}$ , поскольку вблизи рассеивающего объема  $l_{\perp} \sim \lambda$  (см. § 27).

Иные закономерности наблюдаются в *дальней* зоне, для которой, согласно (28.27), время корреляции  $\tau_{\kappa}$  равно времени  $\tau_1 \sim L/\bar{v}$  пребывания отдельной неоднородности в рассеивающем объеме. В дальней зоне показатели экспоненты в формуле (28.25) мал по сравнению с единицей, а  $\Phi_{\pm}(q) \approx \Phi_{\pm}(q_0)$ , так что корреляционная функция поля дается выражением

$$\psi_0(\tau) = e^{-i(\omega + q_0\bar{v})\tau} 8\pi^3 \Phi_{\pm}(q_0) \int_{-\infty}^{\infty} \mathcal{P}_+ \mathcal{P}_- d^3R. \quad (28.31)$$

Для вычисления интеграла необходима, как и в (28.28), конкретизация вида функции  $\mathcal{P}(R)$ . В предельно идеализированном случае, когда функция  $\mathcal{P}(R) = \mathcal{P}_0 = \text{const}$  внутри куба с ребром  $L$  и равна нулю вне этого куба, а скорость дрейфа  $\bar{v}$  перпендикулярна к одной из граней куба, имеем

$$\int_{-\infty}^{\infty} \mathcal{P}_+ \mathcal{P}_- d^3R = \begin{cases} |\mathcal{P}_0|^2 L^3 \left(1 - \frac{|\bar{v}\tau|}{L}\right) & \text{при } |\tau| \leq L/|\bar{v}|, \\ 0 & \text{при } |\tau| > L/|\bar{v}|. \end{cases}$$

Более сложным закономерностям подчинено рассеяние на блуждающих неоднородностях, когда поле скоростей состоит из общей скорости дрейфа  $\bar{v}$  и случайной компоненты  $\tilde{v}$ , т. е.

$v = \bar{v} + \tilde{v}$ . Некоторые аспекты этого случая рассмотрены в [1] и [3—6].

5. Флуктуации интенсивности. В § 27 уже было отмечено, что однократно рассеянное поле распределено по гауссову закону с нулевым средним значением. Для такого поля функция корреляции интенсивности

$$\psi_I(\tau) = \langle I_v(t) I_v(t-\tau) \rangle - \bar{I}_v(t) \bar{I}_v(t-\tau), \quad I_v = |v|^2,$$

выражается через квадрат модуля корреляционной функции поля (см. задачу 12 к гл. I):

$$\psi_I(\tau) = |\psi_v(\tau)|^2. \quad (28.32)$$

Отсюда следует, что характерное время пульсаций интенсивности несколько меньше (в 1,5—2 раза, в зависимости от вида  $\psi_v(\tau)$ ), чем время корреляции самого сигнала  $v(t)$ . Соответственно частотный спектр флуктуаций интенсивности

$$g_I(\Omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\psi_v(\tau)|^2 e^{i\Omega\tau} d\tau = \int_{-\infty}^{\infty} g_v(\Omega_1) g_v(\Omega_1 - \Omega) d\Omega_1 \quad (28.33)$$

несколько шире, чем  $g_v(\Omega)$ .

## § 29. Рассеяние импульсных и модулированных сигналов

Рассеяние монохроматических волн обладает рядом интересных особенностей. Мы ограничимся частным, но важным случаем *квазимонохроматических* сигналов, которые часто используются и в оптике, и в акустике, и в радиофизике. Представив такой сигнал в виде произведения медленно меняющейся (вообще говоря, комплексной) амплитуды  $\mathfrak{A}(t)$  на  $e^{-i\omega t}$  (несущая частота  $\omega$ ), мы должны, очевидно записать сигнал на выходе антенны аналогично формуле (25.46), выведенной для монохроматической волны, но с заменой постоянной амплитуды  $A$  на переменную амплитуду  $\mathfrak{A}(t - b'/c)$  с запаздывающим аргументом ( $b' = |\mathbf{r}' - \mathbf{r}_0| + |\mathbf{r} - \mathbf{r}'|$ ). Если восстановить еще временной фактор  $e^{-i\omega t}$ , опущенный в (25.46), то для квазимонохроматического сигнала получим

$$v(t) = \frac{k_0^2 e^{-i\omega t}}{4\pi} \int_{-\infty}^{\infty} M(\mathbf{r}') \frac{f_0(n'_1) f_1(n'_2)}{|\mathbf{r} - \mathbf{r}'| |\mathbf{r}' - \mathbf{r}_0|} \mathfrak{A}\left(t - \frac{b'}{c}\right) e^{ik_0 b'} \tilde{e}(\mathbf{r}') d^3 r'. \quad (29.1)$$

Для упрощения последующих выкладок введем безразмерную амплитуду  $a(t) = \mathfrak{A}(t)/\mathfrak{A}_m$ , где  $\mathfrak{A}_m = \max |\mathfrak{A}(t)|$ , а через  $\mathcal{S}(\mathbf{r}')$  обозначим произведение неосциллирующих множителей, отли-

чающееся от (25.47) заменой  $A$  на  $\mathfrak{M}$ . В этих обозначениях (29.1) запишется в виде

$$v(t) = e^{-i\omega t} \int_{-\infty}^{\infty} \mathcal{P}(\mathbf{r}') a\left(t - \frac{b'}{c}\right) e^{ik_0 b'} \bar{\epsilon}(\mathbf{r}') d^3 r', \quad (29.2)$$

откуда для временной функции корреляции следует выражение

$$\begin{aligned} \psi_v(t_1, t_2) &= \langle v(t_1) v^*(t_2) \rangle = \\ &= e^{-i\omega \tau} \iint_{-\infty}^{\infty} \mathcal{P}(\mathbf{r}') \mathcal{P}^*(\mathbf{r}'') a\left(t_1 - \frac{b'}{c}\right) a^*\left(t_2 - \frac{b''}{c}\right) e^{ik_0(b' - b'')} \times \\ &\quad \times \psi_e(\tau, \mathbf{r}' - \mathbf{r}'') d^3 r' d^3 r'', \end{aligned} \quad (29.3)$$

где  $\tau = t_1 - t_2$ . Для общности здесь учтена возможная зависимость флуктуаций  $\epsilon$  от времени, а именно введен временной аргумент  $\tau$  в функцию корреляции  $\psi_e$ . Процесс  $v(t)$ , конечно, нестационарен,  $\psi_v$  зависит порознь от  $t_1$  и  $t_2$ . Оценим интервал корреляции  $v(t)$  в случае одиночного импульса (периодическая последовательность импульсов рассмотрена в задаче 5).

Пусть  $a(t)$  — прямоугольный импульс длительности  $T$  ( $a(t) = 1$  при  $|t| \leq T/2$  и равно нулю при  $|t| > T/2$ ). Ясно, что при  $T$ , превышающем время корреляции неоднородностей  $\tau_e$ , время корреляции  $v(t)$  совпадает с  $\tau_e$ . В противоположном же предельном случае  $T \ll \tau_e$  время корреляции будет порядка  $T$ , поскольку произведение  $a(t_1 - b'/c) a^*(t_2 - b''/c)$  обращается в нуль при  $|t_1 - t_2| > T$  для любых значений  $\mathbf{r}'$  и  $\mathbf{r}''$ .

Более детальный анализ выражения (29.3) возможен в предположении, что пространственная длина импульса  $cT$  велика по сравнению с радиусом корреляции неоднородностей  $l_e$ . В этом случае можно выполнить интегрирование по разностному аргументу  $\rho = \mathbf{r}' - \mathbf{r}''$ , как это неоднократно делалось в предыдущих параграфах. В результате

$$\begin{aligned} \psi_v(t_1, t_2) &= \\ &= 8\pi^3 e^{-i\omega \tau} \int_{-\infty}^{\infty} |\mathcal{P}(\mathbf{R})|^2 a\left(t_1 - \frac{B}{c}\right) a^*\left(t_2 - \frac{B}{c}\right) \Phi_e(\tau, \mathbf{q}) d^3 R, \end{aligned} \quad (29.4)$$

где  $B = |\mathbf{r} - \mathbf{R}| + |\mathbf{R} - \mathbf{r}_0|$ .

В отличие от случая монохроматической первичной волны, средняя интенсивность  $\bar{I}_v = \langle |v|^2 \rangle$  зависит теперь от времени:

$$\bar{I}_v(t) = \psi_v(t, t) = 8\pi^3 \int_{-\infty}^{\infty} |\mathcal{P}(\mathbf{R})|^2 \left| a\left(t - \frac{B}{c}\right) \right|^2 \Phi_e(0, \mathbf{q}) d^3 R. \quad (29.5)$$

Величина  $\bar{I}_v$  меньше, чем в случае монохроматического сигнала той же амплитуды, поскольку под знаком интеграла появился множитель  $|a|^2 \leq 1$ . Этот множитель отличен от нуля внутри эллипсоидального слоя, для внутренних точек которого имеем  $|t - B/c| \leq T/2$ , или в развернутой форме

$$c(t - T/2) \leq B(R) \equiv |r - R| + |R - r_0| \leq c(t + T/2). \quad (29.6)$$

Фокусы ограничивающих слой эллипсоидов (рис. 31) совпадают с точками наблюдения ( $r$ ) и излучения ( $r_0$ ), а размеры эллипсоидов с течением времени увеличиваются. Эллипсоидальный слой (29.6) можно назвать *импульсным объемом*. Область, принадлежащая в данный момент одновременно импульсному объему (29.6) и эффективно рассеивающей области  $V_{\text{эфф}}$ , как раз и определяет величину интенсивности  $\bar{I}_v$ .

Толщина импульсного объема  $\delta$  наиболее просто выражается при обратном рассеянии, когда  $r = r_0$  и эллипсоидальный слой (29.6) вырождается в сферический. Как ясно из (29.6),  $\delta = cT/2$ , т. е. толщина импульсного объема равна половине пространственной длины импульса. В общем случае выражения для  $\delta$  довольно сложны, но для не очень длинных импульсов ( $cT \ll |r - r_0|$ ) толщина  $\delta$  определяется простой формулой [1]:

$$\delta = cT/2 \sin(\theta/2), \quad (29.7)$$

где  $\theta$  — угол рассеяния (рис. 31). При рассеянии назад ( $\theta = \pi$ ) из (29.7) получается прежний результат  $\delta = cT/2$ . Отметим, что эллипсоид  $|t - B/c| = \text{const}$  расширяется со скоростью по нормали  $v_n = c/2 \sin(\theta/2)$ .

Временной ход  $\bar{I}_v(t)$  существенно зависит от соотношения между толщиной импульсного объема  $\delta$  и поперечником рассеивающей области  $L$ . В случае *длинного* импульса ( $\delta \geq L$  или  $T \geq (2L/c) \sin(\theta/2)$ ) происходит монотонное увеличение  $\bar{I}_v(t)$ , пока весь рассеивающий объем не окажется внутри импульсного объема (29.6). Время нарастания  $\bar{I}_v(t)$  до максимального значения  $\bar{I}_{\text{max}}$  по порядку величины равно  $L/v_n = (L/c) 2 \sin(\theta/2)$ . В течение такого же времени происходит уменьшение  $\bar{I}_v(t)$  до нуля при выходе из рассеивающей области заднего фронта импульсного объема. Между стадиями нарастания и убывания интен-

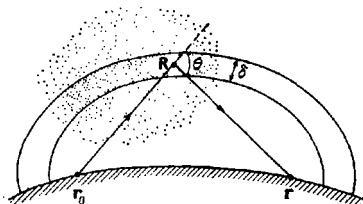


Рис. 31.

сивность  $\bar{I}_v(t)$  постоянна и равна средней интенсивности монохроматического сигнала бесконечной длительности. Если

$$\frac{2L}{c} \sin \frac{\theta}{2} \ll T, \quad (29.8)$$

т. е. времена нарастания и спадания малы по сравнению с длительностью сигнала  $T$ , то рассеянный импульс практически без искажений повторяет форму излученного импульса.

Неравенство (29.8) можно представить и в другом виде, если ввести ширину полосы частот сигнала

$$\Delta\omega \approx \frac{2\pi}{T} \ll \frac{\pi c}{L \sin(\theta/2)}. \quad (29.9)$$

Это условие ограничивает ширину полосы пропускания каналов связи, использующих рассеяние. При передаче и приеме микроволновых сигналов при помощи остронаправленных антенн, когда размеры рассеивающего объема атмосферы  $L$  порядка нескольких десятков километров, полоса пропускания  $\Delta\omega$  оказывается достаточной для передачи телевизионных сигналов ( $\Delta f = \Delta\omega/2\pi \sim 10^6$  Гц).

В противоположном предельном случае  $\delta \ll L$  или  $T \ll (2L/c) \sin(\theta/2)$  (короткий импульс) «время входа» импульса в рассеивающую область, а также «время выхода» из нее имеют порядок длительности импульса:  $\delta/v_n \sim T$ . Время же, в течение которого импульсный объем находится внутри рассеивающей области, приближенно равно  $L/v_n \approx (2L/c) \sin(\theta/2)$ . Для короткого импульса это время значительно превышает длительность сигнала  $T$ , т. е. короткий первичный импульс при рассеянии существенно растягивается — в общем случае примерно в  $(2L/cT) \sin(\theta/2)$  раз, а при обратном рассеянии — в  $2L/cT$  раз.

Следует отметить, что, несмотря на существенное растягивание рассеянного сигнала по сравнению с первичным импульсом, время корреляции по-прежнему остается величиной порядка  $T$ . Это вытекает как из изложенных выше общих соображений, так и из выражения

$$\psi_v(t_1, t_2) = \bar{I}_v \left( \frac{t_1 + t_2}{2} \right) K_{\pi}(\tau) e^{-i\omega\tau} \quad (\tau = t_1 - t_2), \quad (29.10)$$

которое, как можно показать при помощи (29.4), связывает в случае коротких импульсов функцию корреляции поля со средней интенсивностью  $\bar{I}_v$  и «коэффициентом корреляции»



огibaющей первичного импульса

$$K_{\mathbf{n}}(\tau) = \frac{\int_{-\infty}^{\infty} \mathfrak{A}(t) \mathfrak{A}^*(t-\tau) dt}{\int_{-\infty}^{\infty} |\mathfrak{A}(t)|^2 dt}. \quad (29.11)$$

Из (29.10) ясно, что в случае коротких импульсов рассеянное поле представляет собой квазистационарный случайный процесс с переменной дисперсией  $\overline{T}_v(t)$  и с коэффициентом корреляции  $K_{\mathbf{n}}(\tau) e^{-i\omega\tau}$ . Коэффициент корреляции огibaющей  $K_{\mathbf{n}}(\tau)$  обращается при  $\tau > T$  в нуль, откуда и следует, что время корреляции рассеянного поля порядка  $T$ .

### § 30. Рассеяние электромагнитных волн

При рассеянии электромагнитных волн (как и при рассеянии поперечных упругих волн) возникают два новых эффекта, которые отсутствуют в скалярной задаче. Во-первых, рассеяние сопровождается изменением поляризации волны, а во-вторых, если речь идет об анизотропной среде, может происходить трансформация одних типов поляризации в другие. Мы рассмотрим однократное рассеяние электромагнитных волн лишь в *изотропной недиспергирующей* среде, в которой могут распространяться волны только одного типа, а именно поперечные электромагнитные волны.

1. Рассеяние монохроматических электромагнитных волн на покоящихся неоднородностях в изотропной среде. В этом случае уравнения Максвелла

$$\operatorname{rot} \mathbf{H} = \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t}, \quad \operatorname{rot} \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{H}}{\partial t} \quad (30.1)$$

нужно решать совместно с материальным уравнением для изотропной недиспергирующей среды

$$\mathbf{D} = \epsilon \mathbf{E}. \quad (30.2)$$

Материальное уравнение (30.2) исключает из рассмотрения не только рассеяние в анизотропных средах (об особенностях такого рассеяния см., например, в [7]), но и рассеяние на *анизотропных флуктуациях* в изотропной среде (так называемые флуктуации анизотропии). Такие флуктуации возникают, например, в вязких жидкостях из-за случайных поворотов молекул, если последние оптически анизотропны (см. [8] и обзор [9]).

Пусть монохроматическая волна распространяется в среде с постоянной средней диэлектрической проницаемостью  $\bar{\epsilon}$  и с *по-  
коящимися* неоднородностями:

$$\epsilon = \bar{\epsilon} + \tilde{\epsilon}(\mathbf{r}), \quad \bar{\epsilon} = \text{const}, \quad \langle \tilde{\epsilon} \rangle = 0. \quad (30.3)$$

Уравнения Максвелла принимают вид ( $k_0 \equiv \omega/c$ )

$$\text{rot } \mathbf{H} + ik_0 \bar{\epsilon} \mathbf{E} = -ik_0 \tilde{\epsilon} \mathbf{E}, \quad \text{rot } \mathbf{E} - ik_0 \mathbf{H} = 0, \quad (30.4)$$

а их решение может быть представлено при слабых флуктуациях  $\langle (\tilde{\epsilon})^2 \rangle \ll (\bar{\epsilon})^2$  рядами теории возмущений

$$\mathbf{E} = \mathbf{E}_0 + \mathbf{E}_1 + \mathbf{E}_2 + \dots, \quad \mathbf{H} = \mathbf{H}_0 + \mathbf{H}_1 + \mathbf{H}_2 + \dots \quad (30.5)$$

При этом первичное поле  $\mathbf{E}_0$ ,  $\mathbf{H}_0$  удовлетворяет однородным уравнениям

$$\text{rot } \mathbf{H}_0 + ik_0 \bar{\epsilon} \mathbf{E}_0 = 0, \quad \text{rot } \mathbf{E}_0 - ik_0 \mathbf{H}_0 = 0, \quad (30.6)$$

а однократно рассеянное поле — неоднородным уравнениям

$$\text{rot } \mathbf{H}_1 + ik_0 \bar{\epsilon} \mathbf{E}_1 = -ik_0 \tilde{\epsilon} \mathbf{E}_0, \quad \text{rot } \mathbf{E}_1 - ik_0 \mathbf{H}_1 = 0. \quad (30.7)$$

Уравнения для последующих приближений получаются из (30.7) при последовательных заменах  $\mathbf{E}_m \rightarrow \mathbf{E}_{m+1}$ ,  $\mathbf{H}_m \rightarrow \mathbf{H}_{m+1}$ .

Рассеянные поля  $\mathbf{E}_1$  и  $\mathbf{H}_1$  можно найти, используя известные функции Грина, т. е. решение задачи о возбуждении электромагнитных волн точечным электрическим источником  $\mathbf{j}_e = \mathbf{l}_1 \delta(\mathbf{r} - \mathbf{r}')$ . В волновой зоне, т. е. при  $R = |\mathbf{r} - \mathbf{r}'| \gg \lambda$ , элементарный источник создает поля

$$\begin{aligned} \mathbf{E}_{0e}(\mathbf{r}, \mathbf{r}') &= \frac{ik_0^2}{\omega} [\mathbf{n}'_1 [\mathbf{l}_1, \mathbf{n}'_1]] \frac{e^{ikR}}{R}, \\ \mathbf{H}_{0e}(\mathbf{r}, \mathbf{r}') &= \frac{ik_0^2}{\omega} [\mathbf{n}'_1 \mathbf{l}_1] \frac{e^{ikR}}{R}, \end{aligned} \quad (30.8)$$

где  $\mathbf{n}'_1 = (\mathbf{r} - \mathbf{r}')/|\mathbf{r} - \mathbf{r}'| = \mathbf{R}/R$  — единичный вектор, направленный из  $\mathbf{r}'$  в  $\mathbf{r}$ .

Используя (30.8), решение неоднородных уравнений (30.7) можно представить в виде

$$\begin{aligned} \mathbf{E}_1(\mathbf{r}) &= \frac{k_0^2}{4\pi} \int_{-\infty}^{\infty} M(\mathbf{r}') \frac{e^{ikR}}{R} \tilde{\epsilon}(\mathbf{r}') [\mathbf{n}'_1 [\mathbf{E}_0(\mathbf{r}') \mathbf{n}'_1]] d^3r', \\ \mathbf{H}_1(\mathbf{r}) &= \frac{k_0^2 \sqrt{\bar{\epsilon}}}{4\pi} \int_{-\infty}^{\infty} M(\mathbf{r}') \frac{e^{ikR}}{R} \tilde{\epsilon}(\mathbf{r}') [\mathbf{n}'_1 \mathbf{E}_0(\mathbf{r}')] d^3r'. \end{aligned} \quad (30.9)$$

Так как существенно новое по сравнению со скалярной задачей касается только поляризации рассеянного поля, рассмотрим

простейший случай, когда точка наблюдения расположена в зоне Фраунгофера рассеивающего объема ( $r \gg kL^2$ ), а первичное поле представляет собой плоскую волну:

$$E_0(r) = A e^{ik_l r}, \quad k_l = kn_l - k_0 \sqrt{\epsilon} n_l. \quad (30.10)$$

Здесь  $e$  — вектор поляризации первичной волны, вообще говоря — комплексный. Будем считать его нормированным к единице условием  $ee^* = 1$ . В силу поперечности волн, распространяющихся в изотропной среде, вектор поляризации  $e$  перпендикулярен к направлению распространения  $n_l$ . Простой расчет с использованием разложений вида (25.11) дает для однократно рассеянных полей  $E_1$  и  $H_1$  выражения

$$E_1(r) = \frac{k_0^2 A e^{ikr}}{4\pi r} [n_{s0} [en_{s0}]] \int_{-\infty}^{\infty} M(r') \bar{e}(r') e^{-iq_0 r'} d^3 r',$$

$$H_1(r) = \frac{k_0^2 \sqrt{\epsilon} A e^{ikr}}{4\pi r} [n_{s0} e] \int_{-\infty}^{\infty} M(r') \bar{e}(r') e^{-iq_0 r'} d^3 r', \quad (30.11)$$

где  $n_{s0} = r/r$  — единичный вектор, направленный в точку  $r$  из центра рассеивающей области, а  $q_0 = k(n_{s0} - n_{l0})$  — вектор рассеяния, отвечающий центру этой области. Для упрощения записи мы опустим ниже индекс 0 у векторов  $n_{s0}$ ,  $n_{l0}$  и  $q_0$ .

Легко видеть, что поля  $E_1$  и  $H_1$  отличаются от скалярного поля

$$u_1(r) = \frac{k_0^2 A e^{ikr}}{4\pi r} \int_{-\infty}^{\infty} M(r') \bar{e}(r') e^{-iq_0 r'} d^3 r', \quad (30.12)$$

вычисленного при тех же допущениях, только множителями:

$$E_1(r) = [n_s [en_s]] u_1(r), \quad H_1(r) = \sqrt{\epsilon} [n_s e] u_1(r). \quad (30.13)$$

Следовательно, в приближении однократного рассеяния при детерминированной поляризации  $e$  первичной волны любые средние билинейные величины, составленные из компонент полей  $E_1$  и  $H_1$ , могут быть выражены через функцию корреляции скалярного поля  $\psi_s(r_1, r_2) = \langle u_1(r_1) u_1^*(r_2) \rangle$ . Если же поляризация первичного поля случайна, то усреднения по направлениям  $e$  и по ансамблю неоднородностей  $\epsilon$  производятся раздельно в силу статистической независимости этих величин. Иными словами, и при случайной поляризации первичной волны средние от билинейных комбинаций компонент  $E_1$  и  $H_1$  можно выразить через  $\psi_s$  или  $\bar{I}_1$ . Рассмотрим примеры.

2. Средний вектор Пойтинга и эффективный поперечник рассеяния. Среди билинейных комбинаций, составленных из компонент  $E_1$  и  $H_1$ , важную роль играет вектор Пойтинга

$$\mathcal{S}_1 = \frac{c}{8\pi} \operatorname{Re} [E_1 H_1^*] = \frac{c}{16\pi} ([E_1 H_1^*] + [E_1^* H_1]). \quad (30.14)$$

Учитывая, что  $H_1 = \sqrt{\varepsilon} [n_s E_1]$ , и используя выражение (25.34) для  $\bar{I}_1 = \langle |u_1|^2 \rangle$ , находим

$$\begin{aligned} \mathcal{S}_1 &= \frac{c\sqrt{\varepsilon}}{8\pi} n_s \langle |E_1|^2 \rangle = \frac{c\sqrt{\varepsilon} n_s}{8\pi} | [n_s [en_s]] |^2 \bar{I}_1 = \\ &= \frac{c\sqrt{\varepsilon} n_s}{8\pi} | [n_s [en_s]] |^2 \frac{\pi k_0^4 |A|^2}{2r^3} \Phi_\varepsilon(q). \end{aligned} \quad (30.15)$$

Заметим, что в отличие от исходных формул (30.11), это выражение применимо не только в зоне Фраунгофера  $r \gg kL^2$ , но и на гораздо меньших расстояниях  $r \gg r_{\min} = \min \{L, L/\lambda\}$  (см. (25.38)).

Введем, как и в § 26, эффективный поперечник рассеяния единицы объема в единичный телесный угол  $\sigma = |\bar{\mathcal{S}}_1| r^2 / |\mathcal{S}_0| V$ , где  $|\mathcal{S}_0| = c\sqrt{\varepsilon} |A|^2 / 8\pi$  — модуль плотности потока энергии первичной волны. При помощи (30.15) находим

$$\sigma = \frac{|\bar{\mathcal{S}}_1| r^2}{|\mathcal{S}_0| V} = \gamma \sigma_{\text{ск}} = \gamma \frac{\pi}{2} k_0^4 \Phi_\varepsilon(q), \quad (30.16)$$

где  $\sigma_{\text{ск}}$  — «скалярный» поперечник рассеяния (26.12). Множитель

$$\gamma = | [n_s [en_s]] |^2 = | e - n_s (en_s) |^2 = 1 - (n_s e)^2 \quad (30.17)$$

связан с особенностями рассеяния электромагнитных волн по сравнению со скалярной задачей и может быть назван *поляризационным*. Вычислим  $\gamma$  для некоторых частных случаев.

Для *линейно поляризованной* первичной волны  $e$  — вещественный единичный вектор. Если  $\chi$  — угол между  $e$  и  $n_s$  (рис. 32), то  $\gamma = \sin^2 \chi$ . Таким образом, для линейно поляризованной первичной волны

$$\sigma = \sigma_{\text{ск}} \sin^2 \chi. \quad (30.18)$$

Множитель  $\sin^2 \chi$  обращается в нуль в направлениях, коллинеарных с  $e$ , что связано с дипольным характером рассеяния электромагнитных волн в каждом элементе рассеивающего объема.

Произвольную *эллиптически поляризованную* волну можно представить, как известно, в виде суммы двух линейно поляризованных колебаний, сдвинутых по фазе на  $\pi/2$ . Пусть большая ось эл-

липса поляризации направлена вдоль орта  $\mathbf{e}_1$ , а малая—вдоль орта  $\mathbf{e}_2$ . Векторы  $\mathbf{e}_1$  и  $\mathbf{e}_2$  ортогональны к направлению распространения  $\mathbf{n}_i$ , так что  $\mathbf{e}_1$ ,  $\mathbf{e}_2$  и  $\mathbf{n}_i$  образуют ортогональную связку. При эллиптической поляризации

$$\mathbf{e} = p_1 \mathbf{e}_1 + i p_2 \mathbf{e}_2, \quad (30.19)$$

где  $p_1$  и  $p_2$ —вещественные коэффициенты, удовлетворяющие, в силу принятой нами нормировки  $\mathbf{e} \mathbf{e}^* = 1$ , условию  $p_1^2 + p_2^2 = 1$ .

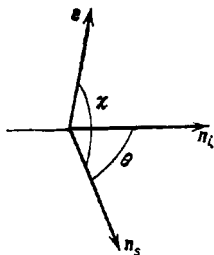


Рис. 32.

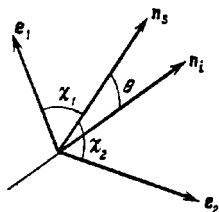


Рис. 33.

Поляризационный множитель  $\gamma$  для волн эллиптической поляризации равен

$$\gamma = 1 - p_1^2 (\mathbf{n}_s \mathbf{e}_1)^2 - p_2^2 (\mathbf{n}_s \mathbf{e}_2)^2 = 1 - p_1^2 \cos^2 \chi_1 - p_2^2 \cos^2 \chi_2, \quad (30.20)$$

где  $\chi_1$  и  $\chi_2$ —углы между вектором  $\mathbf{n}_s$  и осями  $\mathbf{e}_1$  и  $\mathbf{e}_2$  (рис. 33). Подставляя (30.20) в (30.16) и учитывая, что  $p_1^2 + p_2^2 = 1$ , получаем

$$\sigma = (p_1^2 \sin^2 \chi_1 + p_2^2 \sin^2 \chi_2) \sigma_{\text{св}}. \quad (30.21)$$

В отличие от линейно поляризованной первичной волны, сечение (30.21) никогда не обращается в нуль, потому что  $\sin^2 \chi_1$  и  $\sin^2 \chi_2$  не могут обратиться в нуль одновременно. Из (30.13) следует, что в направлениях  $\mathbf{n}_s = \pm \mathbf{e}_1$  и  $\mathbf{n}_s = \pm \mathbf{e}_2$  рассеянное поле поляризовано линейно—в силу уже отмеченного дипольного характера рассеяния. В других направлениях рассеянное поле поляризовано по эллипсу. Параметры эллипса поляризации могут быть найдены из выражений (30.13), если в них подставить (30.19).

В частном случае *круговой поляризации* волны, когда  $p_1 = 1/\sqrt{2}$ ,  $p_2 = \pm 1/\sqrt{2}$  (плюс отвечает правой, а минус—левой поляризацией), находим из (30.20)

$$\gamma = 1 - 1/2 [(\mathbf{n}_s \mathbf{e}_1)^2 + (\mathbf{n}_s \mathbf{e}_2)^2] = 1 - 1/2 (\cos^2 \chi_1 + \cos^2 \chi_2).$$

В силу взаимной ортогональности векторов  $e_1$ ,  $e_2$  и  $n$ , имеем  $\cos^2 \chi_1 + \cos^2 \chi_2 + \cos^2 \theta = 1$ , так что для волн, поляризованных по кругу,

$$\gamma = 1 - 1/2 (1 - \cos^2 \theta) = 1/2 (1 + \cos^2 \theta). \quad (30.22)$$

Таким образом,  $\gamma$  зависит здесь только от угла рассеяния  $\theta$ , т. е. имеет азимутальную симметрию.

Если в точке наблюдения  $r$  измеряется интенсивность рассеянных волн какой-либо одной поляризации, то плотность потока энергии будет, очевидно, меньше плотности полного потока (30.15) и, соответственно, уменьшится величина эффективного сечения рассеяния  $\sigma$ .

Пусть измеряется поток энергии, переносимой компонентой  $E_N = (NE_1)$ , где  $N$  — единичный и, вообще говоря, комплексный вектор, перпендикулярный к направлению рассеяния  $n$ :  $(Nn) = 0$ . Вектор  $N$  описывает поляризационные характеристики приемной антенны в случае радаров или анализатора в случае световых волн. Из (30.13) следует, что

$$E_N = (NE_1) = (N[n_s [en_s]]) u_1,$$

так что роль поляризационного множителя  $\gamma$  играет теперь величина

$$\gamma_N = |(N[n_s [en_s]])|^2 = |(Ne) - (Nn_s)(en_s)|^2 = |(Ne)|^2.$$

Нетрудно убедиться в том, что  $\gamma_N = \gamma$ , если поляризационные характеристики приемной антенны «согласованы» с поляризацией рассеянного излучения, т. е. антенна принимает волну (30.13). В остальных случаях  $\gamma_N < \gamma$ .

3. Рассеяние неполяризованного (естественного) света. С небольшими изменениями полученные результаты переносятся и на *частично поляризованные* поля (ч. I, § 49), когда компоненты первичного поля  $E_0$  случайны. Усреднение по этим величинам можно проводить, как уже было отмечено, независимо от усреднения по реализациям флуктуаций  $\bar{e}$ . Рассмотрим простой, но важный случай *неполяризованной* волны.

Разложим напряженность электрического поля первичной волны на ортогональные компоненты по осям  $e_1$  и  $e_2$ :

$$E_0 = A_1 e_1 + A_2 e_2.$$

В случае неполяризованной волны амплитуды  $A_1$  и  $A_2$  статистически независимы, имеют нулевые средние значения и одинаковые дисперсии:

$$\bar{A}_1 = \bar{A}_2 = 0, \langle A_1 A_2 \rangle = 0, \langle |A_1|^2 \rangle = \langle |A_2|^2 \rangle = 1/2 \langle |E_0|^2 \rangle, \quad (30.23)$$

где  $|E_0|^2 = |A_1|^2 + |A_2|^2$  — полная интенсивность.

Сечение рассеяния неполяризованной волны можно найти из следующих простых соображений. Если бы первичная волна была линейно поляризована вдоль орта  $\mathbf{e}_1$ , то, согласно (30.21), сечение рассеяния равнялось бы  $\sigma = \sigma_1 \equiv \sigma_{\text{св}} \sin^2 \chi_1$  даже при флуктуирующей амплитуде  $A_1$ . При линейной же поляризации вдоль  $\mathbf{e}_2$  мы имели бы  $\sigma = \sigma_2 \equiv \sigma_{\text{св}} \sin^2 \chi_2$ . Поскольку в неполяризованной волне поток энергии делится поровну между двумя ортогональными компонентами, полное сечение рассеяния должно быть равно полусумме  $\sigma_1$  и  $\sigma_2$ :

$$\sigma_{\text{св}} = 1/2 (\sigma_1 + \sigma_2) = 1/2 (\sin^2 \chi_1 + \sin^2 \chi_2) \sigma_{\text{св}}.$$

Но коэффициент при  $\sigma_{\text{св}}$  здесь такой же, как и в случае рассеяния волны с круговой поляризацией, и, следовательно, он равен  $(1 + \cos^2 \theta)/2$  (см. (30.22)). Таким образом,

$$\sigma_{\text{св}} = \frac{1 + \cos^2 \theta}{2} \sigma_{\text{св}}. \quad (30.24)$$

Множитель  $g(\theta) = (1 + \cos^2 \theta)/2$  называют *релевской индикатрисой рассеяния*. Уменьшение  $g(\theta)$  до значения  $1/2$  при  $\theta = \pm \pi/2$  объясняется тем, что в этом направлении вклад в рассеянное поле дает только одна компонента первичной волны, т. е. рассеянное поле оказывается линейно поляризованным.

Выражение для степени поляризации  $P(\theta)$  рассеянного поля можно получить по формуле (49.18) из ч. I:

$$P = \sqrt{1 - \frac{4 \text{Det } J}{(J_{11} + J_{22})^2}}, \quad (30.25)$$

где  $J_{\alpha\beta} = \langle E_\alpha E_\beta^* \rangle$  — элементы поляризационной матрицы в плоскости, перпендикулярной к направлению распространения. Простые вычисления дают принадлежащее Релею выражение

$$P(\theta) = \sqrt{1 - \frac{4 \cos^2 \theta}{(1 + \cos^2 \theta)^2}} = \frac{1 - \cos^2 \theta}{1 + \cos^2 \theta} = \frac{\sin^2 \theta}{1 + \cos^2 \theta}. \quad (30.26)$$

Из (30.26) видно, что  $P=0$  при  $\theta=0$  или  $\pi$ , т. е. поле  $\mathbf{E}_1$  сохраняет естественную поляризацию при рассеянии вперед и назад. При рассеянии под прямым углом ( $\theta = \pm \pi/2$ ) степень поляризации равна единице, а в других направлениях рассеянное поле поляризовано частично ( $0 < P < 1$ ).

4. Частотный спектр рассеянного поля. Некогерентное рассеяние электромагнитных волн в изотропной плазме. При неоднородностях  $\epsilon$ , зависящих от времени, отличия рассеяния электромагнитных волн от аналогичной скалярной задачи тоже могут быть учтены введением в результаты скалярной теории поляризационного множителя  $\gamma = 1 - (\mathbf{n}, \mathbf{e})^2$ . Так, например, в отсутствие регулярного дрейфа

неоднородностей и при  $r \gg r_{\min}$  частотный спектр рассеянного поля

$$g_E(\Omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \langle E_1(t) E_1^*(t-\tau) \rangle e^{i\Omega\tau} d\tau \quad (30.27)$$

в случае сферической первичной волны может быть получен умножением спектра (28.10) на  $\gamma$ :

$$g_E(\Omega) = \gamma \frac{\pi k_0^4 |A|^2 V_{\text{эфф}}}{2r^2 r_0^2} G_s(\omega - \Omega, \mathbf{q}), \quad (30.28)$$

где  $G_s(\omega, \mathbf{k})$  —  $\omega\mathbf{k}$ -плотность флуктуаций диэлектрической проницаемости.

Применим формулу (30.28) к так называемому *некогерентному рассеянию* электромагнитных волн в плазме. Первоначально предполагалось, что рассеяние волн в плазме происходит на свободных электронах, находящихся в тепловом движении, благодаря чему интенсивности полей, рассеянных отдельными электронами, складываются некогерентно (отсюда и возник термин «некогерентное рассеяние»). В дальнейшем выяснилось, однако, что существенную роль в этом явлении играют коллективные процессы<sup>1)</sup>. Мы ограничимся вычислением интенсивности высокочастотного поля, рассеянного в изотропной плазме, опираясь на макроскопическую теорию тепловых флуктуаций (гл. III). В § 31 мы вернемся к рассмотрению некогерентного рассеяния в рамках теории рассеяния на отдельных частицах.

В макроскопической электродинамике свойства изотропной холодной плазмы описываются на высоких частотах ( $\omega \gg \omega_e$ ) диэлектрической проницаемостью

$$\epsilon = 1 - \frac{\omega_e^2}{\omega^2} = 1 - \frac{4\pi e^2 N}{m\omega^2},$$

где  $e$ ,  $m$  и  $N$  — соответственно заряд, масса и концентрация электронов. Таким образом,

$$\bar{\epsilon} = 1 - \frac{4\pi e^2 \bar{N}}{m\omega^2}, \quad \tilde{\epsilon} = \epsilon - \bar{\epsilon} = -\frac{4\pi e^2 \tilde{N}}{m\omega^2}, \quad (30.29)$$

откуда следует связь между  $\omega\mathbf{k}$ -плотностями  $\tilde{\epsilon}$  и  $\tilde{N}$ :

$$G_s(\omega, \mathbf{k}) = \frac{16\pi^2 e^4}{m^2 \omega^4} G_N(\omega, \mathbf{k}). \quad (30.30)$$

<sup>1)</sup> Возможность наблюдать некогерентное рассеяние радиоволн в ионосферной плазме была указана У. Е. Гордоном в 1958 г., а первые наблюдения этого явления осуществил в том же году К. Л. Боулс (см. [10]). Анализ коллективных процессов в плазме посвящены многие работы (см., например, [11]).



Подставляя это выражение для  $G_E$  в формулу (30.28), получаем

$$g_E(\Omega) = 8\pi^2 r_e^2 \gamma \frac{|A|^2 V_{\text{эфф}}}{r^2} G_N(\Omega - \omega, \mathbf{q}), \quad (30.31)$$

где  $r_e = e^2/mc^2$  — классический радиус электрона. Спектр флуктуаций электронной концентрации  $G_N$ , как было показано в задаче 7 гл. III, пропорционален температуре  $T$  и средней концентрации электронов  $\bar{N}$ . Таким образом, спектр рассеянного поля  $g_E$  пропорционален полному числу электронов  $N_E = \bar{N}V$ , находящихся в объеме  $V$ .

Из (30.31) следует, что спектр рассеянного поля  $g_E$  сосредоточен в окрестности частоты первичной волны  $\omega$ , а его форма полностью определяется видом спектра флуктуаций  $G_N(\omega, \mathbf{q})$ . Поэтому частотный спектр рассеянного сигнала  $g_E(\Omega)$  позволяет судить о свойствах и состоянии плазмы. Основанный на этом метод диагностики широко применяется для анализа не только ионосферной, но и лабораторной плазмы («метод некогерентного рассеяния»).

### § 31. Рассеяние на дискретных вкраплениях

1. Поле, рассеянное отдельной частицей. Наряду с рассеянием на объемных неоднородностях, большой интерес представляет и рассеяние на совокупности большого числа тел (частиц, вкраплений и т. д.), случайно расположенных и (или) случайно ориентированных в пространстве. Это очень общая задача, имеющая широкие приложения во многих областях физики.

Строго говоря, проблема рассеяния на системе многих тел относится к статистической схеме 3) (§ 8). Но если дискретные рассеиватели имеют сравнительно небольшие размеры и расположены достаточно редко, то дифракционная задача может быть решена в приближении однократного рассеяния, которое основано на допущении, что каждое вкрапление рассеивает падающую волну так, как если бы других вкраплений не было.

Приближение однократного рассеяния сводит задачу для дискретных рассеивателей, как и в случае непрерывных неоднородностей, к схеме 1), т. е. к излучению полей случайными источниками. В данном случае речь идет о совокупности дискретных источников, статистика которых задана статистикой положений (и ориентаций) рассеивателей. Как мы увидим, «внешние» закономерности однократного рассеяния на дискретных вкраплениях и на непрерывных неоднородностях среды (пространственная корреляция, частотный спектр и т. д.) во многом сходны.

Рассмотрим и здесь упрощенную постановку задачи, а именно рассеяние монохроматической волны на совокупности непод-

вижных частиц, размеры которых малы по сравнению с длиной волны, причем точка наблюдения находится в зоне Фраунгофера объема  $V$ , занятого частицами.

Для малых частиц рассеянное поле можно рассчитывать в дипольном приближении. Пусть  $\mathbf{p}$  — электрический дипольный момент частицы, возникающий под действием электрического поля  $\mathbf{E}$ :

$$\mathbf{p} = \hat{\alpha} \mathbf{E}, \quad (31.1)$$

где  $\hat{\alpha}$  — тензор поляризуемости частицы. Для частиц, малых по сравнению с длиной волны,  $\hat{\alpha}$  можно вычислить, считая поле  $\mathbf{E}$  статистически однородным. Сводка формул для некоторых симметричных тел приведена, например, в монографии [12], посвященной рассеянию волн на малых частицах.

Для сферических частиц из изотропного материала, а также для частиц произвольной формы, но с диэлектрической проницаемостью, близкой к единице, поляризуемость  $\alpha$  является скаляром:

$$\mathbf{p} = \alpha \mathbf{E}. \quad (31.2)$$

В качестве электрического поля в формуле (31.2) следует брать так называемое действующее поле  $\mathbf{E}_d$ , которое отличается от поля первичной волны  $\mathbf{E}_0$  тем, что включает также совокупное поле соседних рассеивателей. В приближении однократного рассеяния (борновское приближение)  $\mathbf{E}_d$  практически не отличается от поля первичной волны  $\mathbf{E}_0$ , которое для простоты мы будем считать полем плоской волны  $\mathbf{E}_0 = A e^{i\mathbf{k}_i \mathbf{r} - i\omega t}$ . Тогда

$$\mathbf{p} = \alpha \mathbf{E}_d \approx \alpha \mathbf{E}_0 = \alpha A e^{i\mathbf{k}_i \mathbf{r}' - i\omega t}, \quad (31.3)$$

где  $\mathbf{r}'$  — радиус-вектор частицы, а  $\mathbf{e}$  — вектор поляризации первичной волны.

Волновое поле, излучаемое в вакууме отдельным диполем с моментом (31.3), может быть рассчитано по формуле (30.8), если подставить туда  $-i\omega \mathbf{p}$  вместо  $\mathbf{I}_1$  (множитель  $e^{-i\omega t}$ , как обычно, опускаем):

$$\mathbf{E}_1(\mathbf{r}) = k_0^2 [\mathbf{n}'_s [\mathbf{e} \mathbf{n}'_s]] \alpha A \frac{\exp(i\mathbf{k} R - i\mathbf{k}_i \mathbf{r}')}{R}. \quad (31.4)$$

Здесь  $R = |\mathbf{r} - \mathbf{r}'|$ ,  $\mathbf{n}'_s = (\mathbf{r} - \mathbf{r}') / |\mathbf{r} - \mathbf{r}'|$  — единичный вектор, направленный из точки  $\mathbf{r}'$  в точку наблюдения  $\mathbf{r}$ . Магнитное поле рассеянной волны в волновой зоне дается выражением

$$\mathbf{H}_1 = [\mathbf{n}'_s \mathbf{E}_1] = [\mathbf{n}'_s \mathbf{e}] k_0^2 \alpha A \frac{\exp(i\mathbf{k} R - i\mathbf{k}_i \mathbf{r}')}{R}. \quad (31.5)$$

2. Среднее значение однократно рассеянного поля. Пусть в объеме  $V$  находятся  $N$  одинаковых частиц с центрами в точках  $\mathbf{r}' = \mathbf{r}_m$  ( $m = 1, 2, \dots, N$ ). В приближении одно-

кратного рассеяния поле в точке наблюдения  $\mathbf{r}$  представляет собой сумму полей, рассеянных отдельными частицами:

$$\mathbf{E}_s(\mathbf{r}) = k_0^2 \sum_{m=1}^N [\mathbf{n}_m [\mathbf{e}\mathbf{n}_m]] \alpha A \frac{\exp(ikR_m + ik_i\mathbf{r}_m)}{R_m}, \quad (31.6)$$

где  $R_m = |\mathbf{r} - \mathbf{r}_m|$ , а  $\mathbf{n}_m = (\mathbf{r} - \mathbf{r}_m)/|\mathbf{r} - \mathbf{r}_m|$  — единичный вектор, направленный от  $m$ -й частицы в точку наблюдения  $\mathbf{r}$ . При учете двукратного рассеяния в качестве  $\mathbf{E}_s$  надо взять сумму первичного поля  $\mathbf{E}_0$  и однократно рассеянного поля (31.6) и т. д.

Заметим, что при учете двукратного рассеяния пренебрежение в действующем поле  $\mathbf{E}_s$  квазистационарным полем индуцированных диполей, ближайших к точке  $\mathbf{r}$ , т. е. пренебрежение полем  $\mathbf{E}_{\text{кв}} \sim \alpha \mathbf{E}_0 / |\mathbf{r} - \mathbf{r}'|^3$  по сравнению с  $\mathbf{E}_0$ , налагает определенное ограничение на среднее расстояние между частицами  $\bar{\rho}$ . Из условия  $\mathbf{E}_{\text{кв}} \ll \mathbf{E}_0$  следует неравенство  $\alpha / \bar{\rho}^3 \ll 1$ , или поскольку  $\bar{n}\bar{\rho}^3 \approx 1$ , где  $\bar{n}$  — средняя концентрация частиц, неравенство

$$\bar{n}\alpha \ll 1. \quad (31.7)$$

При выполнении этого условия возможно как  $\bar{\rho} > \lambda$ , так и  $\bar{\rho} < \lambda$ . Облако частиц, удовлетворяющее неравенству (31.7), можно называть *разреженным* (конечно, в электродинамическом смысле), в отличие от конденсированной среды, в которой  $\bar{n}\alpha \gg 1$ .

Поместим начало координат в центр рассеивающего объема  $V \sim L^3$ , заполненного частицами. Для точек наблюдения  $\mathbf{r}$ , вынесенных в зону Фраунгсфера объема  $V$  ( $r \gg kL^2$ ), можно использовать разложение (25.11):

$$|\mathbf{r} - \mathbf{r}_m| \approx r - \mathbf{n}_s \mathbf{r}_m, \quad \mathbf{n}_m = \frac{\mathbf{r} - \mathbf{r}_m}{|\mathbf{r} - \mathbf{r}_m|} \approx \frac{\mathbf{r}}{r} \equiv \mathbf{n}_s,$$

подставив которое в (31.6), получим

$$\mathbf{E}_s = \mathbf{E}_1 \sum_{m=1}^N e^{-i\mathbf{q}\mathbf{r}_m}, \quad (31.8)$$

где  $\mathbf{q} = k\mathbf{n}_s - \mathbf{k}_i = k(\mathbf{n}_s - \mathbf{n}_i)$  — вектор рассеяния, а

$$\mathbf{E}_1 = k_0^2 [\mathbf{n}_s [\mathbf{e}\mathbf{n}_s]] \alpha A \frac{e^{ikr}}{r} \quad (31.8a)$$

— поле, рассеянное одной частицей, помещенной в начале координат, т. е. в центр области  $V$ .

Для нахождения статистических моментов однократно рассеянного поля нужно задаться законом распределения случайных координат  $\mathbf{r}_m = (x_m, y_m, z_m)$ , а если частицы движутся, то и

законом распределения скоростей  $\mathbf{v}_m = (v_{mx}, v_{my}, v_{mz})$ . Для расчета среднего значения и функции корреляции рассеянного поля необходимы только одночастичная ( $w_1$ ) и бинарная ( $w_2$ ) плотности вероятностей. Примем, что все частицы и все их пары равноправны, т. е. функции  $w_1(\mathbf{r}_m)$  и  $w_2(\mathbf{r}_m, \mathbf{r}_l)$  не зависят от индексов переменных  $\mathbf{r}_m$  и  $\mathbf{r}_l$ .

Усредняя (31.8) по положению частиц  $\mathbf{r}_m$ , получаем среднее поле, рассеянное  $N$  частицами:

$$\bar{E}_s(\mathbf{r}) = \mathcal{E}_1 \left\langle \sum_{m=1}^N e^{-i\mathbf{q}\mathbf{r}_m} \right\rangle = \mathcal{E}_1 N \int_V e^{-i\mathbf{q}\mathbf{r}} w_1(\mathbf{r}) d^3r = \mathcal{E}_1 N f_1(\mathbf{q}), \quad (31.9)$$

где

$$f_1(\mathbf{q}) = \int_V e^{-i\mathbf{q}\mathbf{r}} w_1(\mathbf{r}) d^3r \quad (31.10)$$

— характеристическая функции одночастичного распределения  $w_1(\mathbf{r})$ .

Интенсивность среднего поля  $|\bar{E}_s|^2$ , т. е. когерентная составляющая интенсивности, равна

$$I_{\text{ког}} = |\bar{E}_s|^2 = |\mathcal{E}_1|^2 N^2 |f_1(\mathbf{q})|^2. \quad (31.11)$$

Когерентная составляющая отличается от интенсивности  $|\mathcal{E}_1|^2$  поля, рассеянного одной частицей, множителем  $N^2 |f_1(\mathbf{q})|^2$ , т. е.  $I_{\text{ког}}$  пропорциональна квадрату числа частиц. Величину  $N |f_1(\mathbf{q})|^2$  можно назвать эффективным числом когерентно рассеивающих частиц:

$$N_{\text{ког}} = N |f_1(\mathbf{q})|^2 = \left| \left\langle \sum_{m=1}^N e^{-i\mathbf{q}\mathbf{r}_m} \right\rangle \right|^2. \quad (31.12)$$

При рассеянии вперед ( $\theta = 0$ ,  $q = 0$ ), когда  $f_1(\mathbf{q}) = f_1(0) = 1$ , когерентно излучают все частицы ( $N_{\text{ког}} = N$ ) и при этом

$$(\bar{E}_s)_{\text{вп}} = (N \mathcal{E}_1)_{\text{вп}} = N k_0^2 e \alpha \frac{A}{r} e^{i\mathbf{k}\mathbf{r}}. \quad (31.13)$$

При рассеянии под углом  $N_{\text{ког}}$  резко падает, поскольку характеристическая функция  $f_1(\mathbf{q})$  сравнима с единицей только при  $q \ll 2\pi/L$  или, что то же, при углах рассеяния, лежащих внутри конуса  $\theta \ll \lambda/L \ll 1$  (оценки скорости убывания  $N_{\text{ког}}$  с ростом  $\theta$  даны в задаче 7).

3. Средняя интенсивность. Согласно (31.8) средняя интенсивность однократно рассеянного поля  $\bar{I}_s = \langle |E_s|^2 \rangle$  равна

$$\bar{I}_s = |\mathcal{E}_1|^2 \left\langle \sum_{m,l=1}^N e^{-i\mathbf{q}(\mathbf{r}_m - \mathbf{r}_l)} \right\rangle. \quad (31.14)$$

Выделим из двойной суммы, содержащей  $N^2$  членов,  $N$  слагаемых с  $m = l$ , каждое из которых равно единице, а остальные  $N^2 - N$  слагаемых с  $m \neq l$  усредним при помощи бинарной плотности вероятностей  $w_2$ . В результате получим

$$\bar{I}_s = |\mathcal{E}_1|^2 [N + (N^2 - N) \bar{f}_2(q)], \quad (31.15)$$

где

$$\bar{f}_2(q) \equiv f_2(q, -q) = \int_V w_2(r', r'') e^{-iq(r' - r'')} d^3r' d^3r'', \quad (31.16)$$

$f_2(q_1, q_2)$  — бинарная характеристическая функция. Фактически  $\bar{f}_1(q)$  — это характеристическая функция распределения *относительных* координат. В самом деле, если заменить в (31.16)  $w_2(r', r'') d^3r' d^3r''$  на  $W_2(\rho, R) d^3\rho d^3R$ , где  $\rho = r' - r''$ ,  $R = (r' + r'')/2$ , и выделить интегрирование по  $R$ , то

$$\bar{f}_2(q) = \int W(\rho) \exp(-iq\rho) d^3\rho, \quad W(\rho) = \int W_2(\rho, R) d^3R.$$

Вычитая из (31.15) интенсивность когерентной составляющей поля (31.11), получаем среднюю интенсивность *некогерентного* рассеяния

$$\begin{aligned} \bar{I}_{\text{неког}} &= \bar{I}_s - I_{\text{ког}} = \langle |E_s|^2 \rangle - |\bar{E}_s|^2 = \\ &= |\mathcal{E}_1|^2 \{N[1 - |\bar{f}_1(q)|^2] + (N^2 - N)[\bar{f}_2(q) - |\bar{f}_1(q)|^2]\}. \end{aligned} \quad (31.17)$$

Если положения частиц статистически независимы, т. е.  $w_2(r', r'') = w_1(r') w_1(r'')$ , то  $\bar{f}_2(q) - |\bar{f}_1(q)|^2 = 0$  и второе слагаемое в (31.17), пропорциональное  $N^2 - N$ , обращается в нуль. Следовательно, слагаемое, пропорциональное  $N^2 - N$ , описывает рассеяние на частицах, положения которых *коррелированы* между собой, и его можно рассматривать как характеристику *коллективных* эффектов при рассеянии:

$$\bar{I}_{\text{кол}} = |\mathcal{E}_1|^2 (N^2 - N) [\bar{f}_2(q) - |\bar{f}_1(q)|^2]. \quad (31.18)$$

Первое же слагаемое в (31.17) описывает рассеяние на *независимых* частицах:

$$\bar{I}_{\text{нез}} = |\mathcal{E}_1|^2 N [1 - |\bar{f}_1(q)|^2]. \quad (31.19)$$

Так как  $\bar{f}_1(0) = \bar{f}_2(0) = 1$ , при рассеянии вперед ( $\theta = 0$ ) оба слагаемых в (31.17) обращаются в нуль,  $\bar{I}_{\text{неког}}(0) = 0$ . Но это имеет место только в приближении однократного рассеяния. Можно показать, что уже при учете двукратного рассеяния  $\bar{I}_{\text{неког}}(0) \neq 0$ .

Исключив из рассмотрения рассеяние вперед, т. е. в узкий конус с раствором  $\theta \lesssim \lambda/L$ , можно пренебречь в формуле (31.19)

слагаемым  $|f_1(\mathbf{q})|^2 \ll 1$  (см. задачу 7), что дает

$$\bar{I}_{\text{нез}} \approx |\mathbf{E}_1|^2 N, \quad (31.20)$$

т. е. сумму интенсивностей волн, рассеиваемых независимыми частицами. Если, как чаще всего бывает в приложениях теории,  $N \gg 1$ , то в (31.18) можно пренебречь  $N$  по сравнению с  $N^2$ , и тогда

$$\bar{I}_{\text{кол}} \approx |\mathbf{E}_1|^2 N^2 [\bar{f}_2(\mathbf{q}) - |f_1(\mathbf{q})|^2] \equiv |\mathbf{E}_1|^2 N^2 \nu(\mathbf{q}). \quad (31.21)$$

Представив бинарную функцию распределения  $w_2$  в виде

$$w_2(\mathbf{r}', \mathbf{r}'') = w_1(\mathbf{r}') w_1(\mathbf{r}'') [1 + \mu(\mathbf{r}', \mathbf{r}')], \quad (31.22)$$

можно связать разность  $\nu = \bar{f}_2 - |f_1|^2$  с величиной  $\mu$ , которая обращается в нуль для частиц с независимыми координатами  $\mathbf{r}'$  и  $\mathbf{r}''$ :

$$\begin{aligned} \nu(\mathbf{q}) &= \iint [w_2(\mathbf{r}', \mathbf{r}'') - w_1(\mathbf{r}') w_1(\mathbf{r}'')] e^{-i\mathbf{q}(\mathbf{r}' - \mathbf{r}'')} d^3\mathbf{r}' d^3\mathbf{r}'' = \\ &= \iint w_1(\mathbf{r}') w_1(\mathbf{r}'') \mu(\mathbf{r}', \mathbf{r}'') e^{-i\mathbf{q}(\mathbf{r}' - \mathbf{r}'')} d^3\mathbf{r}' d^3\mathbf{r}''. \end{aligned} \quad (31.23)$$

Если радиус корреляции  $l_\mu$ , т. е. характерный масштаб изменения  $\mu$  по разностной переменной  $\rho = \mathbf{r}' - \mathbf{r}''$ , мал по сравнению с поперечником облака  $L$  (одновременно  $L$  является масштабом изменения  $w_1(\mathbf{r})$  и  $\mu$  по аргументу  $\mathbf{R} = (\mathbf{r}' + \mathbf{r}'')/2$ ), то в пределах сферы  $\rho = |\mathbf{r}' - \mathbf{r}''| \ll l_\mu$  произведение  $w_1(\mathbf{r}') w_1(\mathbf{r}'')$  можно приближенно заменить на  $w_1^2(\mathbf{R})$ , после чего в (31.23) можно выполнить интегрирование по  $\rho$ . Это дает

$$\nu(\mathbf{q}) = 8\pi^3 \int w_1^2(\mathbf{R}) \Phi_\mu(\mathbf{q}, \mathbf{R}) d^3\mathbf{R}, \quad (31.24)$$

где

$$\Phi_\mu(\mathbf{q}, \mathbf{R}) = \frac{1}{8\pi^3} \int \mu(\rho, \mathbf{R}) \exp(-i\mathbf{q}\rho) d^3\rho \quad (31.25)$$

— преобразование Фурье  $\mu(\rho, \mathbf{R})$  по аргументу  $\rho$ . Таким образом,

$$\bar{I}_{\text{кол}} = N^2 |\mathbf{E}_1|^2 \nu(\mathbf{q}) = 8\pi^3 |\mathbf{E}_1|^2 N^2 \int w_1^2(\mathbf{R}) \Phi_\mu(\mathbf{q}, \mathbf{R}) d^3\mathbf{R}. \quad (31.26)$$

4. Средний вектор Пойнтинга. Эффективный поперечник рассеяния. Поскольку в зоне Фраунгофера

$$\mathcal{S}_s = \frac{c}{8\pi} \operatorname{Re} [\mathbf{E}_s \mathbf{H}_s^*] = \frac{c}{8\pi} \mathbf{n}_s |\mathbf{E}_s|^2 = \frac{c}{8\pi} \mathbf{n}_s I_s, \quad (31.27)$$

для среднего значения вектора Пойнтинга  $\bar{\mathcal{S}}_{\text{неког}}$ , отвечающего флуктуационной (некогерентной) компоненте поля, имеем

$$\bar{\mathcal{S}}_{\text{неког}} = \frac{c}{8\pi} \mathbf{n}_s \bar{I}_{\text{неког}} = \frac{c}{8\pi} \mathbf{n}_s (\bar{I}_{\text{нез}} + \bar{I}_{\text{кол}}). \quad (31.28)$$

Обозначим через  $\sigma_1$  поперечник рассеяния отдельной частицы в единичный телесный угол. Согласно (31.8а)

$$\sigma_1 = \frac{|\mathcal{S}_1| r^2}{|\mathcal{S}_0|} = \frac{|\mathbf{g}_1|^2 r^2}{|\mathbf{E}_0|^2} = k_0^4 |[\mathbf{n}_s [\mathbf{e}\mathbf{n}_s]]|^2 \alpha^2 = \gamma k_0^4 \alpha^2, \quad (31.29)$$

где  $\mathcal{S}_0 = \frac{c}{8\pi} \operatorname{Re} [\mathbf{E}_0 \mathbf{H}_0^*]$  — вектор Пойнтинга первичного поля,  $\mathcal{S}_1$  — поток энергии поля (31.4), рассеянного одной частицей, а  $\gamma$  — поляризационный множитель (30.17). Подставляя в (31.28) выражения (31.20) и (31.26) и учитывая (31.29), получаем для суммарного поперечника рассеяния  $\sigma_N$  системы  $N$  частиц в единичный телесный угол выражение

$$\begin{aligned} \sigma_N &= \frac{|\mathcal{S}_{\text{пекор}}| r^2}{|\mathcal{S}_0|} = [N + N^2 \gamma] \sigma_1 = \\ &= \sigma_1 \left[ N + 8\pi^3 N^2 \int \omega_1^2(\mathbf{R}) \Phi_\mu(\mathbf{q}, \mathbf{R}) d^3 R \right]. \end{aligned} \quad (31.30)$$

Вводя среднюю концентрацию частиц

$$\bar{n}(\mathbf{r}) = N \omega_1(\mathbf{r}) \quad (31.31)$$

и учитывая, что  $N = \int \bar{n}(\mathbf{r}) d^3 r$ , можно представить выражение (31.30) в виде

$$\sigma_N = \int \sigma_1 [\bar{n} + 8\pi^3 (\bar{n})^2 \Phi_\mu(\mathbf{q}, \mathbf{R})] d^3 R; \quad (31.32)$$

при этом подынтегральное выражение интерпретируется как сечение рассеяния единичного объема:

$$\sigma = \sigma_{\text{нез}} + \sigma_{\text{кол}}, \quad (31.33)$$

где первое слагаемое

$$\sigma_{\text{нез}} = \sigma_1 \bar{n} \quad (31.34)$$

отвечает рассеянию в пренебрежении корреляцией между положениями частиц, а второе слагаемое

$$\sigma_{\text{кол}} = \sigma_1 \cdot 8\pi^3 (\bar{n})^2 \Phi_\mu(\mathbf{q}, \mathbf{R}) \quad (31.35)$$

— рассеянию с учетом попарной корреляции положений частиц.

В зависимости от конкретных условий соотношение между этими слагаемыми может меняться в широких пределах. Ясно, например, что при малой средней концентрации  $\bar{n}$  преобладает слагаемое  $\sigma_{\text{нез}}$ , а с ростом  $\bar{n}$  — слагаемое  $\sigma_{\text{кол}}$ ; при этом оказывается, что рассеяние на системе частиц с коррелированными положениями может быть описано как рассеяние на макроскопических (объемных) неоднородностях (см. ниже, раздел 5). При рассеянии электромагнитных волн в квазинейтральной плазме

коллективное рассеяние и рассеяние на независимых частицах могут быть сравнимы по величине.

5. Переход к сплошной среде. При макроскопическом рассмотрении диэлектрическая проницаемость разреженного (в смысле неравенства (31.7)) облака частиц равна

$$\varepsilon = 1 + 4\pi n\alpha, \quad (31.36)$$

где  $\alpha$  — поляризуемость частицы, а  $n$  — макроскопическая концентрация частиц, являющаяся, вообще говоря, случайной величиной. Тем самым,

$$\varepsilon = \bar{\varepsilon} + \tilde{\varepsilon}, \quad \bar{\varepsilon} = 1 + 4\pi \bar{n}\alpha, \quad \tilde{\varepsilon} = 4\pi \tilde{n}\alpha. \quad (31.37)$$

Постараемся установить, каковы те статистические характеристики  $\varepsilon$  (или  $n$ ), при которых макроскопическое описание рассеяния, т. е. описание на языке флуктуаций  $\varepsilon$ , дает те же результаты, что и микроскопическое рассмотрение. Говоря о статистических характеристиках  $\varepsilon$ , мы будем иметь в виду только моменты  $\bar{\varepsilon}$  и  $\psi_{\varepsilon}$ .

В силу (31.31) в качестве среднего значения  $\bar{\varepsilon}(\mathbf{r})$  следует, очевидно, взять

$$\bar{\varepsilon}(\mathbf{r}) = 1 + 4\pi \alpha N \bar{w}_1(\mathbf{r}) = 1 + 4\pi \alpha \bar{n}(\mathbf{r}). \quad (31.38)$$

Среднее поле, рассчитанное при помощи (31.38) в борновском приближении для зоны Фраунгофера, совпадает с (31.9).

Для нахождения  $\psi_{\varepsilon}$  отождествим сечение (31.33), вычисленное для системы частиц, с сечением (30.16), описывающим рассеяние на непрерывных объемных неоднородностях. Приравнявая эти два выражения, учитывая (31.29) и сокращая на  $\gamma \pi k_0^2/2$ , получаем

$$\Phi_{\varepsilon}(\mathbf{q}, \mathbf{R}) = \frac{2}{\pi} \bar{n}(\mathbf{R}) \alpha^2 + [4\pi \alpha \bar{n}(\mathbf{R})]^2 \Phi_{\mu}(\mathbf{q}, \mathbf{R}). \quad (31.39)$$

Корреляционная функция  $\psi_{\varepsilon}(\rho, \mathbf{R})$  получается отсюда преобразованием Фурье по  $\mathbf{q}$ :

$$\begin{aligned} \psi_{\varepsilon}(\rho, \mathbf{R}) &= \int \Phi_{\varepsilon}(\mathbf{q}, \mathbf{R}) \exp(i\mathbf{q}\rho) d^3q = \\ &= \frac{2}{\pi} \bar{n}\alpha^2 \int \exp(i\mathbf{q}\rho) d^3q + (4\pi \alpha \bar{n})^2 \int \Phi_{\mu} \exp(i\mathbf{q}\rho) d^3q. \end{aligned}$$

Интеграл в первом слагаемом правой части равен  $8\pi^3 \delta(\rho)$ , а интеграл во втором члене, согласно (31.25), равен  $\mu(\rho, \mathbf{R})$ :

$$\psi_{\varepsilon}(\rho, \mathbf{R}) = (4\pi\alpha)^2 [\bar{n}(\mathbf{R}) \delta(\rho) + (\bar{n})^2(\mathbf{R}) \mu(\rho, \mathbf{R})]. \quad (31.40)$$

Учитывая, что  $\tilde{\varepsilon} = 4\pi \tilde{n}\alpha$ , получаем из (31.40) корреляционную функцию концентрации частиц:

$$\psi_n(\rho, \mathbf{R}) = \bar{n}(\mathbf{R}) \delta(\rho) + (\bar{n})^2(\mathbf{R}) \mu(\rho, \mathbf{R}). \quad (31.41)$$



Первые слагаемые в (31.40) и в (31.41) связаны с дискретностью рассеивателей. Дельта-функция  $\delta(\rho)$  появляется в результате замены реальных частиц точечными диполями. При конечных размерах частиц вместо  $\delta(\rho)$  появилась бы дельтообразная функция, спадающая практически до нуля на расстояниях порядка размера частиц. Второе слагаемое, обусловленное корреляцией положений, напротив, опусывает особенности рассеяния на *коллективе* частиц. При достаточно большой концентрации частиц  $\bar{n}^1$  это слагаемое становится основным, и тогда можно рассматривать рассеяние на системе частиц с коррелированными положениями как рассеяние на объемных макроскопических неоднородностях сплошной среды с корреляционной функцией

$$\varphi_e = (4\pi\alpha\bar{n})^2 \mu(\rho, R).$$

Условие, при котором можно при *любых* углах рассеяния пренебречь «дискретным» слагаемым  $\sigma_{\text{нез}}$  по сравнению с членом  $\sigma_{\text{кол}}$ , имеет вид

$$8\pi^2\bar{n}(\Phi_\mu)_{\min} \gg 1, \quad (31.42)$$

где  $(\Phi_\mu)_{\min}$  — минимальное (по  $q$ ) значение спектра. При выполнении же неравенства  $8\pi^2\bar{n}(\Phi_\mu)_{\max} \ll 1$  можно пренебречь коллективными эффектами. Если ни одно из этих неравенств не выполняется во всем диапазоне изменения угла рассеяния  $0 < \theta < \pi$ , то надо пользоваться либо полным сечением (31.33), либо, определив из условия  $8\pi^2\bar{n}\Phi_\mu \sim 1$  граничный угол  $\theta_{\text{гр}}$ , пользоваться различными предельными формулами при  $\theta > \theta_{\text{гр}}$  и  $\theta < \theta_{\text{гр}}$ .

6. О некогерентном рассеянии в плазме. Этот вопрос мы уже затрагивали в § 30. Здесь мы рассмотрим рассеяние на свободных электронах в плазме под другим углом зрения, а именно как пример явления, в котором, в силу особенностей парной корреляции, коллективное рассеяние оказывается пропорциональным не второй, а *первой* степени концентрации электронов  $N_e$ .

Двумерная плотность вероятностей  $w_2(\mathbf{r}_1, \mathbf{r}_2)$  в случае плазмы, находящейся в равновесии при температуре  $T$ , выражается распределением Больцмана:

$$w_2(\mathbf{r}_1, \mathbf{r}_2) = \text{const} \cdot \exp \left[ -\frac{U_{12}(\mathbf{r}_1, \mathbf{r}_2)}{\Theta} \right],$$

где  $\Theta = kT$  — температура, выраженная в энергетических единицах, а  $U_{12}(\mathbf{r}_1, \mathbf{r}_2) = U_{12}(\mathbf{r}_1 - \mathbf{r}_2)$  — эффективная потенциальная

<sup>1)</sup> Совместимой, однако, с условием  $4\pi\alpha\bar{n} \ll 1$ , при котором можно не учитывать вклада квазистационарного поля. Учет  $E_{\text{кн}}$  привел бы к формуле Лорентца — Лоренца [13].

энергия электрона, находящегося в точке  $\mathbf{r}_2$ , в поле другого электрона, расположенного в точке  $\mathbf{r}_1$ .

Предположим, что частицы равномерно заполняют большой объем  $V$ , так что  $\omega_1(\mathbf{r}) = \text{const} = 1/V$ , и рассмотрим случай, когда энергия взаимодействия двух электронов  $U_{12}$  мала по сравнению с энергией теплового движения  $\Theta = kT$ . Тогда приближенно двумерная плотность будет равна

$$[\omega_2(\mathbf{r}_1, \mathbf{r}_2) \approx \frac{1}{V^2} \left(1 - \frac{U_{12}}{\Theta}\right)]. \quad (31.43)$$

Сравнивая (31.43) с формулой (31.23), в которой нужно положить  $\omega_1 = 1/V$ , получаем для функции  $\mu$ , характеризующей корреляцию положений электронов, приближенное выражение

$$\mu(\mathbf{r}_1 - \mathbf{r}_2) \approx -\frac{U_{12}}{\Theta} = -\frac{U(\mathbf{r}_1 - \mathbf{r}_2)}{\Theta}.$$

В случае изотропной плазмы  $U_{12} = e\phi_{12}$ , где  $\phi_{12}$  — дебаевский потенциал, отличающийся от кулоновского потенциала  $e/|\mathbf{r}_1 - \mathbf{r}_2|$  экспоненциальным множителем  $\exp\{-|\mathbf{r}_1 - \mathbf{r}_2|/d\}$ , где  $d = (\Theta/8\pi e^2 \bar{N}_e)^{1/2}$  — дебаевский радиус (радиус экранирования) [14], а  $\bar{N}_e$  — средняя концентрация электронов. Таким образом,

$$\mu(\mathbf{r}_1 - \mathbf{r}_2) \approx -\frac{e^2}{\Theta |\mathbf{r}_1 - \mathbf{r}_2|} e^{-|\mathbf{r}_1 - \mathbf{r}_2|/d}.$$

Преобразование Фурье этой функции, равное

$$\Phi_\mu(\mathbf{q}) = -\frac{1}{8\pi^3 \bar{N}_e} \frac{1}{2 + q^2 d^2},$$

обратно пропорционально концентрации  $\bar{N}_e$ , так что для поперечника коллективного рассеяния по формуле (31.35) находим

$$\sigma_{\text{кол}} = -\frac{\sigma_1 \bar{N}_e}{2 + q^2 d^2}. \quad (31.44)$$

В результате суммарный поперечник рассеяния (31.33) оказывается равным

$$\sigma = \sigma_1 \bar{N}_e - \frac{\sigma_1 \bar{N}_e}{2 + q^2 d^2} = \sigma_1 \bar{N}_e \frac{1 + q^2 d^2}{2 + q^2 d^2},$$

или, учитывая, что поперечник рассеяния отдельной частицы  $\sigma_1 = \gamma k_0^2 \alpha^2$  в случае электрона равен  $\gamma r_e^2 = \gamma e^4 / m^2 c^4$ ,

$$\sigma = \gamma r_e^2 \bar{N}_e \frac{1 + q^2 d^2}{2 + q^2 d^2}. \quad (31.45)$$

В предельном случае коротких волн ( $qd \gg 1$ ) из (31.45) следует, что

$$\sigma = \gamma \bar{N}_e r_e^2 \quad (qd \gg 1). \quad (31.46)$$

Таким образом, суммарный эффект при рассеянии коротких волн на тепловых флуктуациях электронной концентрации оказывается таким же, как при рассеянии на  $N$  свободных электронах. Это связано с независимым характером движения электронов в пределах области  $\Delta r \sim 1/q \ll d$ , целиком лежащей внутри сферы дебаевского радиуса  $d$ .

В противоположном случае, когда длина волны  $\Lambda_q = 2\pi/q$  много больше дебаевского радиуса, движение электронов уже нельзя считать независимым от движения ионов. В этом случае сечение рассеяния будет

$$\sigma = 1/2 \gamma \bar{N}_e r_e^2 \quad (qd \ll 1), \quad (31.47)$$

т. е. вдвое меньше, чем в (31.46).

### Задачи

1. Показать, что полный поперечник рассеяния единичного объема  $\sigma$  в случае мелкомасштабных флуктуаций ( $kl_e \ll 1$ ) равен

$$\sigma_0 = 2\pi^2 k_0^4 \Phi_e(0) = \frac{k_0^4}{4\pi} \int_{-\infty}^{\infty} \psi_e(\rho) d^3\rho, \quad (1)$$

а в случае крупномасштабных флуктуаций ( $kl_e \gg 1$ ) —

$$\sigma_0 = 1/2 \pi k_0^2 \int_{-\infty}^{\infty} \Phi_e(x_1, x_2, 0) dx_1 dx_2 = 1/4 k_0^2 \int_{-\infty}^{\infty} \psi_e(0, 0, \rho_3) d\rho_3. \quad (2)$$

Решение. Формула (1) сразу же следует из определения (26.13), так как при  $kl_e \ll 1$

$$\Phi_e(q) \approx \Phi_e(0) = \frac{1}{8\pi^3} \int_{-\infty}^{\infty} \psi_e(\rho) d^3\rho.$$

В случае же  $kl_e \gg 1$  надо учесть, что существенный вклад в (26.13) дают только малые углы рассеяния  $\theta \ll 1/kl_e \ll 1$ . Вводя переменные интегрирования  $x_1 = k\theta \cos \varphi$ ,  $x_2 = k\theta \sin \varphi$  и учитывая, что при малых углах рассеяния  $\Phi_e(q) \approx \Phi_e(x_1, x_2, 0)$  и  $do = \sin \theta d\theta d\varphi \approx \frac{1}{k^2} dx_1 dx_2$ , получаем формулу (2).

2. Оценить полный поперечник рассеяния  $\sigma_0$  световых волн в случае турбулентных флуктуаций со спектральной плотностью  $\Phi_e(x) = \mathcal{A} C_e^2 (x^2 + x_0^2)^{-11/6}$ , не имеющей особенности при  $x \rightarrow 0$  ( $x_0 = 2\pi/L_0$  — волновое число, отвечающее внешнему масштабу турбулентности  $L_0$ ).

Решение. В оптическом диапазоне длина волны  $\lambda$  мала по сравнению с внутренним масштабом турбулентности  $l_0$ . Поэтому для расчета  $\sigma_0$  можно воспользоваться формулой (2) предыдущей задачи, что дает (при  $\mathcal{A} = 0,033$ )

$$\sigma_0 = 3/5 \pi^2 k^2 \mathcal{A} C_e^2 x_0^{-5/6} \approx 0,0091 k^2 C_e^2 L_0^{5/6}.$$

# I

## ОСНОВЫ СТАТИСТИЧЕСКОЙ РАДИОТЕХНИКИ

### Глава 1 ОБЩИЕ СВЕДЕНИЯ О СИГНАЛАХ

#### 1.1. Основные определения. Классификация сигналов

Любая радиотехническая система имеет своей целью доставить информацию о каком-либо явлении, событии или состоянии некоторой материальной системы в условиях, когда источник информации удален от места ее приема. Это может быть текстовая информация (радиосвязь), информация о координатах или скорости движения удаленного объекта (радиолокация), изображение объекта (телевидение), информация о состоянии удаленной материальной системы (телеметрия), команды на изменение состояния объекта (командные радиолинии) и т.д.

Совокупность символов, несущих информацию о каком-либо событии или состоянии системы, называется сообщением. Сигналом называется физический носитель сообщения, в частности, радиосигналом называется электромагнитное колебание, параметры которого изменяются по закону, отображающему сообщение.

В любой реальной радиотехнической системе параметры электромагнитного колебания, являющегося переносчиком сообщения, подвержены действию случайных помех, под которыми понимаются все сторонние возмущения, приводящие к искажению сигнала. Источниками помех могут служить внешние шумы, шумы приемника, посторонние сигналы, поступающие на вход приемника, искажения сигнала при распространении его через среду и т.п. Источником случайной помехи может служить и сам передаваемый сигнал, поступивший на вход приемника не по основному направлению распространения, а в результате отражения от поверхности земли, ионосферы или каких-либо тел либо неоднородностей (многолучевость). Случайный характер помехи обусловлен в этом случае невозможностью строго учета условий распространения сигнала по побочным лучам.

По характеру действия помехи можно разделить на аддитивные и мультипликативные. Принимаемое колебание при действии аддитивной помехи  $n(t)$  можно представить в виде

$$\xi_u(t) = u(t) + n(t)$$

где  $u(t)$  — передаваемый сигнал. Аддитивную помеху, независимо от ее природы будем называть шумом.

При действии мультипликативной помехи  $v(t)$  принимаемое колебание можно представить в виде

$$\xi_u(t) = v(t) u(t)$$

где  $v(t)$  — случайная неотрицательная функция.

В дальнейшем мы ограничимся рассмотрением случая аддитивной помехи. При этом будем полагать, что помехи могут быть приведены ко входу приемного устройства, и будем рассматривать входное колебание как смесь полезного сигнала и шума.

По степени полноты априорной информации о сигнале различают случайные, детерминированные и квазидетерминированные сигналы.

*Случайные сигналы* характеризуются тем, что их значения в любой момент времени заранее непредсказуемы с требуемой точностью. Случайный сигнал является частным случаем случайного процесса, представляющего изменение какой-либо физической величины по непредсказуемому закону. Вид функции  $x_i(t)$ , описывающей случайный сигнал (случайный процесс), становится известным только после его приема. Так, например, многократно включая генератор шума на некоторое фиксированное время, мы каждый раз будем получать различные по форме отрезки функции  $x_i(t)$ . Конкретные сигналы  $x_i(t)$ , получаемые при каждом включении генератора шума, будем называть *реализациями случайного сигнала*  $\xi(t)$ . Случайные сигналы (процессы) будем обозначать буквами греческого алфавита, а их конкретные реализации — буквами латинского алфавита. Вид каждой реализации  $x_i(t)$  заранее неизвестен. Однако совокупность всех полученных реализаций  $\{x_i(t)\}$  принадлежащих единому источнику, обладающему некоторыми физическими свойствами, будет обладать определенными статистическими характеристиками, являющимися единственной формой задания случайного сигнала.

Сигналы, заранее известные для любого момента времени, описываемые заданной функцией времени  $u(t)$ , называются *детерминированными*.

Разновидностью случайных сигналов являются *квазидетерминированные сигналы*, реализации которых описываются функциями заданного вида, но содержат один или несколько случайных параметров  $\alpha, \beta, \dots$  не зависящих от времени. Реализации квазидетерминированного сигнала могут быть записаны в виде

$$u_i(t) = u(t; \alpha_i, \beta_i, \dots)$$

и полностью определяются значениями случайных параметров  $\alpha_i, \beta_i, \dots$ . Квазидетерминированные сигналы называют также сигналами заданной формы со случайными значениями параметров.

Кроме того, следует выделить часто встречающуюся в статистической радиотехнике разновидность случайных сигналов — сигналы с медленно меняющимися случайными параметрами, реализации которых имеют вид

$$u_i(t) = u[t; \alpha_i(t), \beta_i(t), \dots].$$

В отличие от квазидетерминированных сигналов в этом случае параметры  $\alpha_i(t)$ ,  $\beta_i(t)$ , ... являются не случайными величинами, а случайными функциями времени, изменяющимися, однако, гораздо медленнее, чем сам сигнал  $u_i(t)$ , при фиксированных значениях этих параметров. На коротких отрезках времени, пока можно не считаться с изменениями параметров, сигнал ведет себя как квазидетерминированный. Однако при более протяженных интервалах наблюдения его параметры должны уже рассматриваться как случайные функции времени.

В зависимости от задач, которые ставятся при обработке сигнала, сигнал или его параметры могут выступать:

а) как случайное событие, когда требуется оценить сам факт наличия или отсутствия какого-либо сигнала (задача обнаружения);

б) как случайная величина, когда требуется оценить значение случайного параметра квазидетерминированного сигнала (задача оценки параметров);

в) как случайная функция (случайный процесс) когда требуется оценить реализацию случайного сигнала  $u_i(t)$  или случайно изменяющегося параметра сигнала  $\alpha_i(t)$  (задача фильтрации).

По форме представления зависимости сигнала от времени все сигналы можно подразделить на непрерывные, дискретные и цифровые.

*Непрерывными* называются сигналы, заданные во всех точках временной оси или некоторого ее отрезка. Их реализации  $u_i(t)$ , как правило, непрерывные функции времени (рис. 1.1).

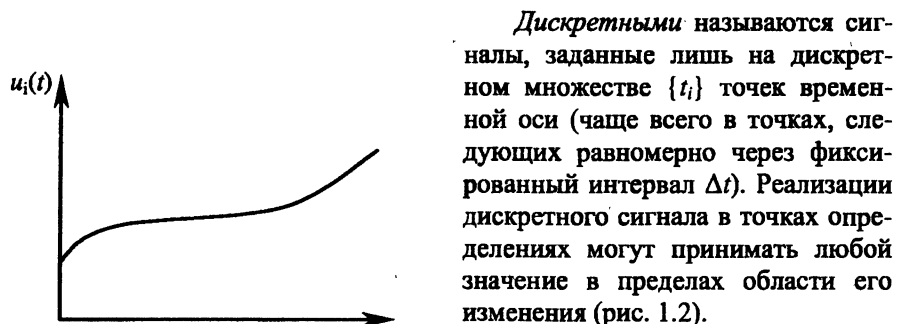


Рис. 1.1

*Дискретными* называются сигналы, заданные лишь на дискретном множестве  $\{t_i\}$  точек временной оси (чаще всего в точках, следующих равномерно через фиксированный интервал  $\Delta t$ ). Реализации дискретного сигнала в точках определения могут принимать любое значение в пределах области его изменения (рис. 1.2).

*Цифровые* сигналы представляют собой частный случай дискретных сигналов, когда реализации  $u_i(t_k)$  в точках определения

сигнала  $t_k$  могут принимать одно из фиксированных дискретных значений, определяемых числами с ограниченным количеством разрядов (рис. 1.3).

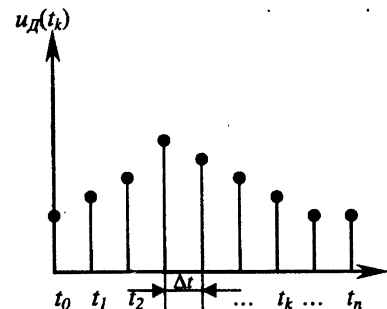


Рис. 1.2

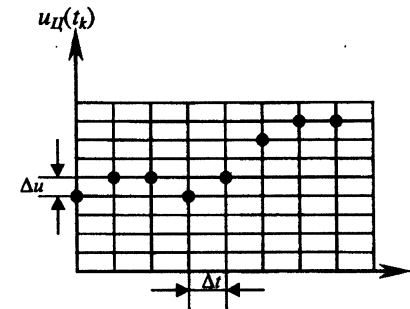


Рис. 1.3

Следует отметить, что в радиотехнических системах мы имеем дело только со случайными и квазидетерминированными сигналами: передача по радиоканалу детерминированных сигналов, полностью известных в месте приема по априорной информации, очевидно, лишена смысла.

## 1.2. Формы представления сигналов

Наиболее естественной формой представления сигналов является задание закона изменения его во времени  $u(t)$ . Однако эта форма представления сигналов не является единственной, а может рассматриваться как одна из разновидностей спектрального представления сигналов [6].

При спектральном представлении сигнал  $u(t)$  задается в виде линейной комбинации определенным образом выбранных сигналов известной формы (базисных функций), определенных на том же временном интервале, на котором определен и сигнал  $u(t)$  (например, на интервале  $[-T/2, T/2]$ ). Набор коэффициентов при базисных функциях, определяющий (при заданной системе базисных функций) сигнал  $u(t)$ , называется его *спектром*.

Если сигнал  $u(t)$  может быть представлен в виде суммы спектральных составляющих, отвечающих счетному множеству базисных функций  $\varphi_i(t)$ :

$$u(t) = \sum_{i=-\infty}^{\infty} a_i \varphi_i(t), \quad (1.1)$$

то спектр такого сигнала, представляемый набором коэффициентов  $a_i$ , называется дискретным. В этом случае сигнал  $u(t)$  и базисные функции  $\varphi_i(t)$

имеют одинаковую размерность, а коэффициенты  $a_i$  соответственно безразмерные. Если для представления сигнала  $u(t)$  недостаточно дискретного набора базисных функций  $\varphi_i(t)$ , а требуется несчетное множество базисных функций  $\varphi(a, t)$ , отличающихся значениями непрерывно изменяющегося параметра  $a$ , то сигнал представляется в виде интеграла:

$$u(t) = \int_{-\infty}^{\infty} S(a) \varphi(a, t) da. \quad (1.2)$$

Спектр такого сигнала характеризуется функцией  $S(a)$  непрерывной переменной  $a$  (непрерывный спектр).

В этом случае роль элементарного безразмерного коэффициента играет величина  $S(a) da$ , величина же  $S(a)$  представляет его плотность по параметру  $a$  и имеет размерность, обратную размерности  $a$ .

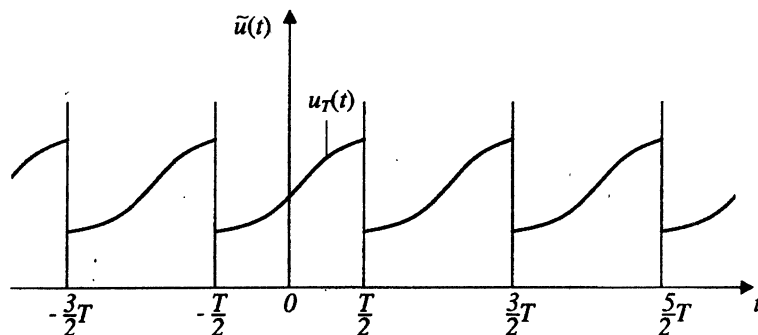


Рис. 1.4

Рассмотрим наиболее распространенную форму спектрального представления сигналов, когда в качестве базисных функций используются гармонические сигналы. Рассмотрим действительный сигнал  $u_T(t)$ , заданный на интервале  $[-T/2, T/2]$ . Дополним сигнал  $u_T(t)$  до периодического сигнала  $\tilde{u}_T(t)$ , повторяя функцию  $u_T(t)$  с периодом  $T$  вне интервала  $[-T/2, T/2]$  (рис. 1.4). Разлагая периодическую функцию  $\tilde{u}_T(t)$  в ряд Фурье [37], получим:

$$u_T(t) = \tilde{u}_T(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos 2\pi f_k t + b_k \sin 2\pi f_k t), \quad (1.3)$$

где

$$a_k = \frac{2}{T} \int_{-T/2}^{T/2} u(t) \cos 2\pi f_k t dt; \quad b_k = \frac{2}{T} \int_{-T/2}^{T/2} u(t) \sin 2\pi f_k t dt; \quad f_k = \frac{k}{T}. \quad (1.4)$$

Используя подстановку

$$a_k = c_k \cos \varphi_k; \quad b_k = c_k \sin \varphi_k, \quad (1.5)$$

разложение (1.3) можно представить в виде

$$u_T(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (c_k \cos (2\pi f_k t - \varphi_k)), \quad |t| \leq \frac{T}{2}, \quad (1.6)$$

где

$$c_k = \sqrt{a_k^2 + b_k^2}; \quad \varphi_k = \arctg (b_k/a_k). \quad (1.7)$$

При проведении математических операций над сигналами вместо тригонометрической формы представления сигналов (1.4), (1.6) удобно использовать представление в комплексной форме с применением экспоненциальных базисных функций  $e^{j2\pi f_k t}$ .

Для этого вводится двусторонний комплексный спектр путем "расщепления" каждой спектральной сопоставляющей разложения (1.6) действительного сигнала  $u(t)$  на две комплексно-сопряженные компоненты для частот  $f_k$  и  $-f_k$ , равные соответственно

$$\dot{C}(f_k) = |\dot{C}(f_k)| e^{j\varphi_k} = \frac{1}{2} c_k \cos \varphi_k - j \frac{1}{2} c_k \sin \varphi_k = \frac{a_k}{2} - j \frac{b_k}{2}; \quad (1.8)$$

$$\dot{C}(-f_k) = |\dot{C}(f_k)| e^{j\varphi_k} = \frac{a_k}{2} + j \frac{b_k}{2}, \quad (1.9)$$

где точка сверху используется для обозначения комплексной функции (числа).

Разложение (1.6) в комплексной форме имеет вид

$$u_T(t) = \sum_{k=-\infty}^{\infty} \dot{C}(f_k) e^{j2\pi f_k t} = \sum_{k=-\infty}^{\infty} \dot{C}(f_k) e^{j2\pi k t/T}, \quad (1.10)$$

где

$$\dot{C}(f_k) = \frac{1}{T} \int_{-T/2}^{T/2} u(t) (\cos 2\pi f_k t - j \sin 2\pi f_k t) dt = \frac{1}{T} \int_{-T/2}^{T/2} u(t) e^{-j2\pi f_k t} dt. \quad (1.11)$$

Введя обозначение  $\Delta f = f_{k+1} - f = 1/T$  и перейдя в (1.10) к пределу при  $T \rightarrow \infty$ , получим разложение в виде интеграла Фурье для сигнала  $u(t)$  неограниченной длительности

$$u(t) = \lim_{T \rightarrow \infty} u_T(t) = \lim_{\substack{T \rightarrow \infty \\ \Delta f \rightarrow 0}} \sum_{k=-\infty}^{\infty} \frac{1}{\Delta f} \dot{C}(f_k) e^{j2\pi f_k t} \Delta f = \\ = \int_{-\infty}^{\infty} \dot{S}(f) e^{j2\pi f t} df, \quad (1.12)$$

где

$$\dot{S}(f) = \lim_{\Delta f \rightarrow 0} \frac{\dot{C}(f)}{\Delta f} = \lim_{\substack{T \rightarrow \infty \\ \Delta f \rightarrow 0}} \frac{1}{T \Delta f} \int_{-T/2}^{T/2} u(t) e^{-j2\pi f t} dt = \\ = \int_{-\infty}^{\infty} u(t) e^{-j2\pi f t} dt. \quad (1.13)$$

Величину  $\dot{S}(f)$ , представляющую спектральную плотность комплексной амплитуды  $\dot{C}(f)$ , называют комплексным спектром. Сигнал  $u(t)$  и комплексный спектр  $\dot{S}(f)$  связаны соотношениями (1.12) и (1.13), которые носят название *преобразований Фурье*. Более распространена запись преобразований Фурье с использованием в качестве аргумента круговой частоты  $\omega = 2\pi f$ :

$$u(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{S}(\omega) e^{j\omega t} d\omega,$$

$$\dot{S}(\omega) = \int_{-\infty}^{\infty} u(t) e^{-j\omega t} dt,$$

где  $\dot{S}(\omega)$  — спектральная плотность комплексной амплитуды на 1 Гц (а не на 1 рад/с). При такой записи появляется нарушающий симметрию формул коэффициент  $1/2\pi$ , связанный с тем, что спектральная плотность на единицу круговой частоты (рад/с) равна  $\dot{S}(\omega)/2\pi$ .

Заметим, что для действительного сигнала комплексной является лишь форма записи. В самом деле, суммируя спектральные составляющие двустороннего комплексного спектра для симметричных частот  $f$  и  $-f$  и учитывая (1.8), (1.9), получаем

$$\dot{S}(f) e^{j2\pi f t} + \dot{S}(f) e^{-j2\pi f t} = |\dot{S}(f)| [e^{j(2\pi f t - \varphi_f)} + e^{-j(2\pi f t - \varphi_f)}] = 2 |\dot{S}(f)| \cos(2\pi f t - \varphi_f),$$

т.е. двустороннему комплексному спектру  $\dot{S}(f)$  соответствует односторонний гармонический спектр со спектральной плотностью амплитуды  $2 |\dot{S}(f)|$  и фазой  $\varphi_f$ .

Легко убедиться, что преобразования Фурье (1.12), (1.13) справедливы и для комплексной функции  $\omega(t) = u(t) + jv(t)$  (примерами такой функции являются комплексный сигнал и комплексный спектр). Применяя преобразования Фурье к действительным функциям  $u(t)$  и  $v(t)$ , получаем

$$u(t) = \int_{-\infty}^{\infty} \dot{S}_u(f) e^{j2\pi f t} df, \quad v(t) = \int_{-\infty}^{\infty} \dot{S}_v(f) e^{j2\pi f t} df,$$

где

$$\dot{S}_u(t) = \int_{-\infty}^{\infty} u(t) e^{-j2\pi f t} dt, \quad \dot{S}_v(t) = \int_{-\infty}^{\infty} v(t) e^{-j2\pi f t} dt.$$

Соответственно получаем

$$\dot{\omega}(t) = u(t) + jv(t) = \int_{-\infty}^{\infty} [\dot{S}_u(f) + j\dot{S}_v(f)] e^{j2\pi f t} df = \int_{-\infty}^{\infty} \dot{S}(f) e^{j2\pi f t} df,$$

где

$$\dot{S}(f) = \dot{S}_u(f) + j\dot{S}_v(f) = \int_{-\infty}^{\infty} [u(t) + jv(t)] e^{-j2\pi f t} dt = \int_{-\infty}^{\infty} \dot{\omega}(t) e^{-j2\pi f t} dt.$$

Полученные выражения по форме записи ничем не отличаются от преобразований Фурье для действительных функций (действительных сигналов). Однако для комплексных сигналов двусторонний спектр уже не будет комплексно-сопряженным и сведение его к одностороннему спектру становится вообще невозможным.

На комплексные сигналы распространяются также и формулы (1.10), (1.11) для представления сигналов ограниченной длительности в виде рядов Фурье.

Для сигнала  $u_T(t)$  отличного от нуля в ограниченном интервале  $[-T/2, T/2]$ , формула (1.13) принимает вид

$$\dot{S}(f) = \int_{-T/2}^{T/2} u(t) e^{j2\pi ft} dt.$$

В этом случае для представления сигнала на интервале  $[-T/2, T/2]$  можно использовать как интеграл Фурье (1.12), так и ряд Фурье (1.10). Эти два представления отличаются лишь вне интервала  $[-T/2, T/2]$ : если интеграл Фурье правильно отображает сигнал  $u_T(t)$  на всей временной оси (вне интервала  $[-T/2, T/2]$  в рассматриваемом случае он дает значения, равные нулю), то ряд Фурье вне интервала  $[-T/2, T/2]$  перестает соответствовать отображаемому сигналу и вместо нулевого значения дает периодическую функцию  $\tilde{u}(t)$  (рис. 1.4).

Определим энергию сигнала  $u(t)$  через его комплексный спектр  $\dot{S}(f)$ . Рассмотрим сначала действительный сигнал  $u_T(t)$  ограниченной длительности. На интервале  $[-T/2, T/2]$  его можно представить в виде ряда (1.10) и получить следующее выражение для его энергии:

$$\begin{aligned} \int_{-T/2}^{T/2} u^2(t) dt &= \int_{-T/2}^{T/2} \left[ \sum_{k=-\infty}^{\infty} \dot{C}(f_k) e^{j\frac{2\pi k t}{T}} \right]^2 dt = \\ &= \int_{-T/2}^{T/2} \left[ \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \dot{C}(f_k) \dot{C}(f_l) e^{j\frac{2\pi(k+l)t}{T}} \right] dt = \\ &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \dot{C}(f_k) \dot{C}(f_l) \int_{-T/2}^{T/2} e^{j\frac{2\pi(k+l)t}{T}} dt. \end{aligned} \quad (1.14)$$

Определим значения интегралов, входящих в (1.14):

$$\begin{aligned} \int_{-T/2}^{T/2} e^{j\frac{2\pi(k+l)t}{T}} dt &= \int_{-T/2}^{T/2} \cos \frac{2\pi(k+l)t}{T} dt + \\ &+ j \int_{-T/2}^{T/2} \sin \frac{2\pi(k+l)t}{T} dt = T \frac{\sin \pi(k+l)}{\pi(k+l)} = \\ &= \begin{cases} 0 & \text{при } k+l \neq 0 \\ T & \text{при } k+l = 0 \end{cases} \end{aligned} \quad (1.15)$$

Учитывая, что  $\dot{C}(-f) = \dot{C}^*(f)$  (звездочка — знак комплексного сопряженного числа), на основании (1.15) получаем:

$$\int_{-T/2}^{T/2} \dot{C}(f_k) e^{j\frac{2\pi k t}{T}} \dot{C}^*(f_l) e^{-j\frac{2\pi l t}{T}} dt = \begin{cases} 0 & \text{при } k \neq l, \\ T \dot{C}(f_k) \dot{C}^*(f_k) & \text{при } k = l. \end{cases} \quad (1.16)$$

Функции  $\dot{\phi}_k(t)$  и  $\dot{\phi}_l(t)$ , удовлетворяющие условию

$$\int_{-T/2}^{T/2} \dot{\phi}(t) \dot{\phi}^*(t) dt = 0,$$

называются ортогональными на интервале  $[-T/2, T/2]$ . Выражение (1.16) свидетельствует об ортогональности на интервале  $T$  существования сигнала  $u_T(t)$  всех его спектральных составляющих. Формула (1.14) с учетом (1.15) принимает вид:

$$\int_{-T/2}^{T/2} u^2(t) dt = \sum_{k=-\infty}^{\infty} T \dot{C}(f_k) \dot{C}^*(f_k) = \sum_{k=-\infty}^{\infty} T |\dot{C}(f_k)|^2. \quad (1.17)$$

Слагаемые  $T |\dot{C}(f_k)|^2$  в (1.17) представляют энергию отдельных спектральных составляющих двустороннего комплексного спектра на интервале  $T$  существования сигнала  $u_T(t)$ , величина  $|\dot{C}(f_k)|^2$  представляет среднюю мощность каждой из этих спектральных составляющих.

Формула (1.17), определяющая энергию сигнала  $u_T(t)$  через энергии его спектральных составляющих, называется *равенством Парсеваля*.

Учитывая, что  $T = 1/\Delta f$  и переходя в (1.17) к пределу, получаем равенство Парсеваля для сигнала, определяемого на неограниченном интервале:



$$\int_{-\infty}^{\infty} u^2_T(t) dt = \lim_{T \rightarrow \infty} \int_{-T/2}^{T/2} u^2_T(t) dt = \lim_{T \rightarrow \infty} \sum_{k=-\infty}^{\infty} \left| \frac{C(f_k)}{\Delta f} \right|^2 \Delta f = \int_{-\infty}^{\infty} |S(f)|^2 df \quad (1.18)$$

Из (1.18) следует, что величина  $|S(f)|^2$  представляет спектральную плотность энергии сигнала (энергетический спектр).

Равенство Парсеваля характеризует тождественность выражений для энергии сигнала, соответствующих временному и частотному его представлениям. Обе эти формы представления сигнала совершенно равноправны. Временное представление — это разложение сигнала с использованием в качестве базисных функций бесконечно узких по оси времени импульсов (с бесконечно протяженным частотным спектром). Частотное представление — это разложение по гармоническим базисным сигналам, бесконечно узким по частотному спектру и бесконечно протяженным по времени. Равенство Парсеваля означает, что при обоих этих формах представления сигнала его полная энергия равна сумме энергий всех спектральных составляющих (сумме энергий бесконечно коротких импульсов при временном представлении и бесконечно узких участков частотного спектра при частотном представлении). Равенство Парсеваля является следствием ортогональности базисных функций на интервале существования сигнала. Ортогональность гармонических базисных функций на интервале  $(-\infty, \infty)$  следует из (1.16), а ортогональность базисных функций в виде бесконечно узких, примыкающих друг к другу импульсов очевидна, так как они не перекрываются по времени.

Гармонический спектр долгое время был единственным применявшимся в радиотехнике. С развитием дискретной техники широкое распространение получили спектральные разложения по дискретным базисным функциям (функция Радемахера, Уолша и др.) [6]. Следует отметить, что и для аналоговых сигналов гармонический спектр является наиболее естественным лишь при рассмотрении систем с постоянными параметрами. Для систем с переменными параметрами более удобными оказываются другие базисные функции [6, 14]. В качестве базисных функций для разложения сигналов могут также использоваться системы ортогональных полиномов (полиномы Лежандра и др.) [6].

В некоторых случаях удобно использовать геометрическую форму представления сигнала, когда базисным функциям ставятся в соответствие оси ортогональной системы координат в многомерном пространстве сигнала, а составляющие спектра представляют собой значения этих координат, задающие сигнал  $u(t)$ . При этом дискретному спектру ставится в соответствие евклидово пространство с числом измерений, равным числу составляющих спектра  $a_i$  (базисных функций  $\phi_i(t)$ ). Сигнал в этом пространстве может задаваться в векторной форме, при этом составляющие спектра определяют компоненты вектора, отображающего сигнал.

Если базисные функции нормированы, т.е.

$$\int_{-T/2}^{T/2} |\phi_i(t)|^2 dt = 1,$$

то квадрат модуля вектора, равный в соответствии с метрикой  $N$ -мерного евклидова пространства

$$R^2_u = \sum_{i=1}^N a_i^2,$$

представляет собой энергию сигнала, а квадрат расстояния между двумя сигналами в евклидовом пространстве равен:

$$R^2_{u_1, u_2} = \sum_{i=1}^N (a_{2i} - a_{1i})^2.$$

Для неограниченного дискретного спектра пространство сигнала — счетное бесконечномерное, и суммы, представляющие квадрат модуля вектора, отображающего сигнал ("модуля сигнала") и квадрат расстояния между сигналами, содержат бесконечное число членов.

Непрерывному спектру, имеющему спектральную плотность  $S(\alpha)$ , соответствует бесконечномерное пространство сигнала с несчетным множеством измерений. Квадрат модуля сигнала и квадрат расстояния между сигналами в этом пространстве выражается не через суммы, а через интегралы:

$$R^2_u = \int_0^{\infty} S^2(\alpha) d\alpha; \quad R^2_{u_1, u_2} = \int_0^{\infty} [S_1(\alpha) - S_2(\alpha)]^2 d\alpha.$$

В частности, при временном представлении дискретных сигналов (дискретный "временной спектр") эти выражения принимают вид

$$R^2_u = \sum_{i=-\infty}^{\infty} u^2(t_i), \quad R^2_{u_1, u_2} = \sum_{i=-\infty}^{\infty} [u_2(t_i) - u_1(t_i)]^2. \quad (1.19)$$

При временном представлении непрерывных сигналов (сплошной "временной спектр")

$$R^2_u = \int_{-\infty}^{\infty} u^2(t) dt; \quad R^2_{u_1, u_2} = \int_{-\infty}^{\infty} [u_2(t) - u_1(t)]^2 dt. \quad (1.20)$$

### 1.3. Дискретизация непрерывных сигналов.

#### Теорема Котельникова

В настоящее время широкое распространение получили дискретные радиоканалы, т.е. каналы, по которым передаются цифровые сигналы. Они широко используются и для передачи непрерывных сообщений. Это объясняется следующими обстоятельствами: 1) большей устойчивостью цифровых сигналов к возможным искажениям в канале; 2) возможностью непосредственного ввода информации для последующей обработки в цифровых ЭВМ; 3) возможностью преобразований непрерывного сигнала в цифровой и обратно практически с любой заданной точностью.

Рассмотрим сначала возможность дискретизации непрерывного сигнала. В.А. Котельниковым было доказано [7], что дискретизация непрерывного сигнала в случае, когда ширина его частотного спектра ограничена, при соответствующем выборе шага квантования по времени теоретически обеспечивает возможность абсолютно точного восстановления непрерывного сигнала. Теорему Котельникова можно сформулировать следующим образом: *непрерывный сигнал  $u(t)$ , спектр которого ограничен сверху частотой  $\mathcal{F}_B$ , полностью определяется своими значениями,*

отсчитываемыми в дискретных точках через интервалы  $\Delta t = 1/2F_0$ , где  $F_0 \geq \mathcal{F}_B$ , при этом значение сигнала в произвольный момент времени  $t$  равно

$$u(t) = \sum_{k=-\infty}^{\infty} u(k\Delta t) \frac{\sin [2\pi F_0 (t - k\Delta t)]}{2\pi F_0 (t - k\Delta t)} \quad (1.21)$$

Будем считать, что сигнал  $u(t)$  может быть представлен интегралом Фурье через комплексный спектр  $\dot{S}(f)$ :

$$u(t) = \int_{-\infty}^{\infty} \dot{S}(f) e^{j2\pi f t} df = \int_{-F_0}^{F_0} \dot{S}(f) e^{j2\pi f t} df, \quad (1.22)$$

так как по условию теоремы Котельникова  $\dot{S}(f) = 0$  при  $|f| > F_0$ .

Комплексный спектр  $\dot{S}(f)$  в области существования  $[-F_0, F_0]$ , может быть отображен рядом Фурье вида (1.10), применимым и для отображения комплексных функций:

$$\dot{S}(f) = \sum_{k=-\infty}^{\infty} \dot{D}_k e^{j2\pi k f / 2F_0}, \quad -F_0 \leq f \leq F_0, \quad (1.23)$$

где коэффициенты разложения  $\dot{D}_k$  согласно (1.11) равны

$$\dot{D}_k = \frac{1}{2F_0} \int_{-F_0}^{F_0} \dot{S}(f) e^{-j2\pi k f / 2F_0} df. \quad (1.24)$$

При этом следует помнить, что ряд (1.23) правильно отображает функцию  $\dot{S}(f)$  только в пределах области  $[-F_0, F_0]$ . Учитывая, что  $1/2F_0 = \Delta t$ , записываем (1.24) в виде

$$\dot{D}_k = \Delta t \int_{-F_0}^{F_0} \dot{S}(f) e^{-j2\pi k f \Delta t} df. \quad (1.25)$$

Сравнивая (1.25) с (1.22), видим, что коэффициенты  $\dot{D}_k$  представляют произведение интервала  $\Delta t$  на значения сигнала  $u(t)$  в точках  $-k\Delta t$ :

$$\dot{D}_k = \Delta t u(-k\Delta t). \quad (1.26)$$

Подставляя (1.26) в (1.23), получаем

$$\dot{S}(f) = \Delta t \sum_{k=-\infty}^{\infty} u(-k\Delta t) e^{j2\pi k f \Delta t} = \Delta t \sum_{k=-\infty}^{\infty} u(k\Delta t) e^{-j2\pi k f \Delta t}. \quad (1.27)$$

Подставляя теперь (1.27) в (1.22) и учитывая, что  $\Delta t = 1/2F_0$ , имеем

$$u(t) = \Delta t \sum_{k=-\infty}^{\infty} u(k\Delta t) \int_{-F_0}^{F_0} e^{j2\pi f (t - k\Delta t)} df = \sum_{k=-\infty}^{\infty} u(k\Delta t) \frac{\sin 2\pi F_0 (t - k\Delta t)}{2\pi F_0 (t - k\Delta t)}. \quad (1.28)$$

Строго говоря, условия теоремы Котельникова для реальных сигналов не удовлетворяются, поскольку они представляют процессы, ограниченные во времени, и поэтому их спектр не ограничен. Однако практически всегда можно ограничить сверху спектр реального сигнала достаточно большой частотой  $\mathcal{F}_B$  так, чтобы составляющие спектра за ее пределами были незначительными. Оценим допускаемую при этом погрешность, полагая, что "усечение" спектра осуществляется идеально, без искажений для частот  $f \leq \mathcal{F}_B$ . Искажение сигнала  $u(t)$  при таком "усечении" его спектра будем характеризовать среднеквадратическим отклонением "усеченного" сигнала от истинного на интервале существования последнего:

$$\sigma_c = \sqrt{\frac{1}{T} \int_0^T \Delta u^2(t) dt}, \quad (1.29)$$

где  $\Delta u(t)$  — отклонение сигнала с "усеченным" спектром от реального;  $T$  — длительность сигнала.

Поскольку реальный и "усеченный" сигналы имеют одинаковый спектр в интервале частот  $[-\mathcal{F}_B, \mathcal{F}_B]$ , спектр разностного сигнала  $\Delta u(t)$  определяется отброшенной частью спектра сигнала  $u(t)$  для  $|f| > \mathcal{F}_B$ . Поэтому, используя равенство Парсеваля (1.18), получаем

$$\int_0^T \Delta u^2(t) dt < \int_{-\infty}^{\infty} \Delta u^2(t) dt = \int_{\mathcal{F}_B}^{\infty} 2 |S(jf)|^2 df = \Delta E, \quad (1.30)$$

где  $\Delta E$  — энергия, приходящаяся на неучитываемую часть спектра реального сигнала, лежащую выше частоты  $\mathcal{F}_B$ .

Из (1.29) и (1.30) следует, что

$$\sigma_c < \sqrt{\Delta E / T}. \quad (1.31)$$

Относительную погрешность, возникающую при "усечении" спектра сигнала, будем характеризовать отношением среднего квадрата искажения сигнала к его средней мощности  $\mathcal{P}_c$ . Тогда получим ее оценку в виде:

$$\frac{\sigma_c^2}{\mathcal{P}_c} < \frac{\Delta E}{T \mathcal{P}_c} = \frac{\Delta E}{E_c}, \quad (1.32)$$

т.е. относительная погрешность меньше относительной энергии  $\Delta E/E_c$ , соответствующей неучитываемой части спектра.

Число дискретных отсчетов, минимально необходимых для задания сигнала длительностью  $T$  при ограничении его спектра частотой  $\mathcal{F}_B$  и  $\mathcal{F}_B T \gg 1$ , равно

$$B = (T/\Delta t) + 1 = T 2\mathcal{F}_B + 1 \approx 2\mathcal{F}_B T. \quad (1.33)$$

При этом сигнал приближенно, с погрешностью, определяемой неравенствами (1.31), (1.32), представляется суммой

$$u(t) \approx \sum_{k=1}^B u(k\Delta t) \frac{\sin 2\pi \mathcal{F}_B (t - k\Delta t)}{2\pi \mathcal{F}_B (t - k\Delta t)}. \quad (1.34)$$

Величину  $B$  называют числом степеней свободы или базой сигнала.

Теорема Котельникова указывает на симметрию свойств временно-го и спектрального (частотного) представлений сигнала: как сигнал, ограниченный во времени, может быть задан дискретным набором его спектральных составляющих (дискретным частотным спектром), так и сигнал с ограниченным спектром может быть задан набором его отсчетов в дискретные моменты времени (дискретным "временным спектром"). Эта симметрия обоих представлений отображается разложениями (1.10) и (1.21).

Теорема Котельникова имеет фундаментальное значение в теории информации, статистической радиотехнике и ряде других прикладных наук. Вместе с тем она находит и прямое применение при решении практических задач (выбор частоты опроса телеметрических систем, определение режимов работы систем с адаптацией и т.д.). Следует отметить, что разложение Котельникова является не единственным и не всегда лучшим способом интерполяции значений непрерывной функции по ее дискретным отсчетам.

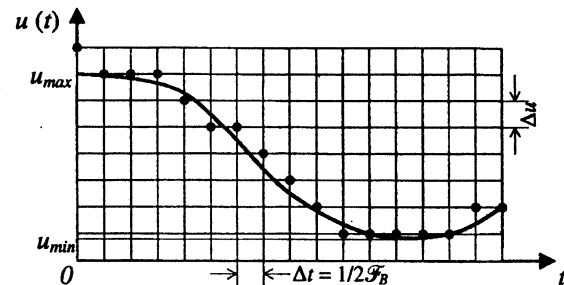


Рис. 1.5

Рассмотрим теперь дискретизацию значений сигнала (квантования сигнала по уровню) для перехода к цифровой форме. Дискретизация значений сводится к замене бесконечного множества возможных значений непрерывного сигнала конечным множеством дискретных его значений. Для этого весь интервал от  $u_{min}$  до  $u_{max}$  возможных значений сигнала разбивается на конечное число дискретных уровней (рис. 1.5) с интервалом между уровнями (шагом квантования)  $\Delta_u$ . Значения  $u(t)$  сигнала в дискретные моменты отсчета  $t_k$  заменяются ближайшим дискретным значением  $u_i$  в соответствии с правилом

$$u_i = \Delta_u \left[ \frac{u(t_k)}{\Delta_u} + \frac{1}{2} \right]_{цч},$$

где

$$\left[ \frac{u(t_k)}{\Delta_u} + \frac{1}{2} \right]_{цч}$$

— целая часть величины, стоящей в скобках.

На Рис. 1.5 дискретные значения сигнала обозначены точками. Легко видеть, что погрешность при таком представлении ("шум квантования") не превышает половины шага квантования  $\Delta_u$ .

При малом  $\Delta_u$ , когда распределение возможных значений сигнала в пределах шага квантования можно считать равномерным, среднеквадратическое значение этой погрешности [16].

$$\sigma_k = \Delta_u / 2\sqrt{3}. \quad (1.35)$$

При заданной величине допустимой ошибки квантования  $\delta u_{\text{кв}} = u_i - u(t_k)$  (заданной величине  $\sigma_k$  и соответственно при заданном шаге квантования  $\Delta u = 2\sqrt{3}\sigma_k$ ) требуемое число уровней квантования

$$n = \left\lceil \frac{u_{\max} - u_{\min}}{\Delta u} \right\rceil_{\text{цч}} + 1. \quad (1.36)$$

Таким образом, непрерывный сигнал  $u(t)$  может быть преобразован в цифровой. Погрешности, вносимые при таком преобразовании, определяемые формулами (1.31) и (1.35), могут быть уменьшены до требуемой величины соответствующим выбором шага квантования по времени и по уровню.

## Глава 2

# СЛУЧАЙНЫЕ СОБЫТИЯ, ВЕЛИЧИНА, ФУНКЦИЯ И ИХ ВЕРОЯТНОСТНЫЕ ХАРАКТЕРИСТИКИ

## 2.1. Случайное событие и вероятность

Под случайными событиями понимаются события, которые могут произойти или не произойти в результате данного опыта. Если при  $N$ -кратном повторении какого-либо опыта относительное число появлений  $i$ -го события по мере увеличения общего числа реализаций этого опыта  $N$  стремится к некоторому пределу

$$P_i = \lim_{N \rightarrow \infty} (n_i/N), \quad (2.1)$$

то говорят о статистической устойчивости результатов данного опыта.

Числа  $P_i$ , позволяющие при этом количественно сравнить события по степени их возможности, называются *вероятностями* этих событий. Чем чаще имеет место данный исход опыта, тем больше его вероятность. Достоверное событие, которое непременно должно произойти при каждом испытании, имеет максимально возможное значение вероятности  $P_i = 1$ . В противоположность ему невозможное событие, т.е. событие, которое в данном опыте не может произойти, имеет минимальное значение вероятности  $P_i = 0$ . Для всех остальных событий, которые могут иметь место при данном опыте, но не при каждом его повторении, вероятность лежит в пределах  $0 < P_i < 1$ .

Для некоторых опытов вероятности отдельных исходов могут быть предсказаны умозрительно на основе анализа условий опыта. Это, как правило, те случаи, когда можно выделить такие элементарные исходы опыта, ни одному из которых нельзя отдать предпочтение, так что все они должны приниматься равновероятными. Так, например, при бросании монеты нет никаких оснований отдать предпочтение одному из двух возможных исходов опыта (падение монеты гербом вверх или вниз) и обоим им должна приписываться одинаковая вероятность, равная  $1/2$ . То же относится и к бросанию игральной кости в виде идеального кубика, когда любая из его граней с одинаковой вероятностью  $P_i = 1/6$  может оказаться

верхней. Нет оснований отдать предпочтение и какой-либо из игральным карт при их сдаче из хорошо перетасованной колоды и т.д.

В более сложных случаях, когда не удастся умозрительно определить вероятности элементарных исходов опыта, единственным методом их определения остается непосредственный подсчет относительной частоты появления отдельных исходов при многократном воспроизведении опыта. В этом случае опыт одновременно позволяет убедиться в статистической устойчивости его результатов и найти вероятности отдельных исходов.

В большинстве случаев, представляющих практический интерес, мы сталкиваемся со сложными событиями, представляющими определенные сочетания исходов ряда элементарных опытов. Прибегать для оценки вероятности такого сложного события каждый раз к опыту чрезвычайно трудно, а иногда и просто невозможно (например, при проектировании, когда объект испытаний вообще отсутствует). Возникает задача косвенной оценки вероятности сложных событий через известные вероятности других событий, логически с ними связанных. Систему таких косвенных методов оценки вероятностей, позволяющих свести к минимуму экспериментальные оценки, дает *теория вероятностей*.

Введем некоторые понятия, являющиеся основополагающими для теории вероятностей.

События называются *несовместными* в данном опыте, если никакие два из них не могут появиться вместе. Если в результате каждой конкретной реализации опыта обязательно должно появиться хотя бы одно из событий  $A_1, A_2, \dots, A_n$ , то такая группа событий называется *полной*. *Противоположными* называются два несовместных события  $A$  и  $\bar{A}$ , образующих полную группу. События называются *независимыми*, если вероятность появления одного из них не зависит от того, имело место или нет другое из этих событий. Если же вероятность события  $A$  зависит от того, имело ли место событие  $B$ , то такие события называются *зависимыми*. В этом случае можно ввести условную вероятность  $P(A|B)$ , представляющую вероятность события  $A$  при условии, что имело место событие  $B$ .

*Суммой* нескольких событий называется событие, заключающееся в появлении хотя бы одного из них. *Произведением* нескольких событий называется событие, заключающееся в совместном появлении всех этих событий. Понятия суммы и произведения употребляются здесь не в арифметическом смысле, а обозначают соответствующие логические операции.

В основе косвенных методов оценки вероятностей лежат теоремы сложения и умножения вероятностей.

*Теорема сложения вероятностей*: вероятность суммы несовместных событий  $A_1, A_2, \dots, A_n$  равна сумме вероятностей этих событий:

$$P(A_1 + A_2 + \dots + A_n) = \sum_{i=1}^n P(A_i). \quad (2.2)$$

Справедливость равенства (2.2) вытекает из того, что в случае несовместных событий число исходов опыта, при которых имеет место любое из событий  $A_1, A_2, \dots, A_n$  будет равно сумме исходов, отвечающих каждому из них. Следует еще раз отметить, что равенство (2.2) справедливо *только для несовместных событий*.

*Теорема умножения вероятностей*: вероятность произведения двух событий равна произведению вероятности одного из них на условную вероятность другого, соответствующую условию, что имело место первое событие:

$$P(AB) = P(A) P(B|A). \quad (2.3)$$

В справедливости (2.3) легко убедиться, сопоставляя соответствующие числа исходов опыта. Действительно, если в  $N$  опытах  $n$  раз произошло событие  $A$  и при этих  $n$  исходах  $k$  раз одновременно имело место и событие  $B$ , то при неограниченном возрастании числа опытов  $N$  и соответственно исходов  $A$  имеем:

$$P(A) = \lim_{N \rightarrow \infty} (n/N), \quad P(B|A) = \lim_{n \rightarrow \infty} (k/n).$$

Поскольку из общего числа  $N$  опытов только в  $k$  случаях имели место одновременно события  $A$  и  $B$ , то получим:

$$P(AB) = \lim_{N \rightarrow \infty} (k/N) = \lim_{N \rightarrow \infty} \left[ \frac{n}{N} \frac{k}{n} \right] = P(A) P(B|A).$$

Вероятность произведения двух событий можно также записать в виде

$$P(AB) = P(B) P(A|B). \quad (2.4)$$

Из (2.3) и (2.4) получаем

$$P(B|A) = \frac{P(B) P(A|B)}{P(A)}. \quad (2.5)$$

В частном случае, если события  $A$  и  $B$  *независимы*,  $P(B|A) = P(B)$  и

$$P(AB) = P(A) P(B). \quad (2.6)$$

Проиллюстрируем применение теорем сложения и умножения вероятностей на простейших примерах.

Найдем вероятность того, что из хорошо перетасованной колоды, содержащей 36 карт, будет вытащена карта данной масти. В данном случае требуется найти вероятность суммы 9 несовместных равновероятных событий, заключающихся в вытаскивании одной из карт данной масти, вероятность каждого из которых равна  $1/36$ . Следовательно, согласно теореме сложения искомая вероятность  $P = 9(1/36) = 1/4$ .

Найдем теперь вероятность сложного события, заключающегося в том, что из урны, содержащей 7 красных и 3 белых шара, неотличимых на ощупь, в двух попытках будут вынуты 1 красный и 1 белый шар.

Рассмотрим сначала случай, когда перед очередной попыткой вынутый ранее шар возвращается в урну. В этом случае результаты обеих попыток представляют независимые события. Поскольку вероятность вынуть любой из 10 шаров в каждой попытке равна  $1/10$ , то в соответствии с теоремой сложения красный шар может быть вынут с вероятностью  $7/10$ , а белый — с вероятностью  $3/10$ . Вероятность вынуть в двух независимых попытках разноцветные шары в любой последовательности

$$\begin{aligned} P_{\text{ркб}} &= P(\text{кб} + \text{бк}) = P(\text{кб}) + P(\text{бк}) = \\ &= P(\text{к})P(\text{б}) + P(\text{б})P(\text{к}) = 0,7 \cdot 0,3 + 0,3 \cdot 0,7 = 0,42. \end{aligned}$$

Здесь обозначения событий  $\text{кб}$  и  $\text{бк}$  соответствуют последовательности, в которой вынимаются красный ( $\text{к}$ ) и белый ( $\text{б}$ ) шары.

В случае, когда шар, вынутый в первой попытке, в урну не возвращается, вероятность того или иного исхода во второй попытке зависит от результатов первой попытки. Если в первой попытке был вынут белый шар, то условные вероятности вынуть при второй попытке из урны, содержащей 7 красных и 2 белых шара, красный или белый шар соответственно равны

$$P(\text{к}|\text{б}) = 7/9, \quad P(\text{б}|\text{к}) = 2/9.$$

Если же в первой попытке был вынут красный шар, то условные вероятности вынуть красный или белый шар во второй попытке соответственно равны

$$P(\text{к}|\text{к}) = 6/9 = 2/3, \quad P(\text{б}|\text{к}) = 3/9 = 1/3.$$

Вероятность вынуть в двух попытках разноцветные шары

$$P_{\text{ркб}} = P(\text{к})P(\text{б}|\text{к}) + P(\text{б})P(\text{к}|\text{б}) = 0,7 \cdot (1/3) + 0,3 \cdot (7/9) = 0,467.$$

Пользуясь теоремами сложения и умножения вероятностей, получим некоторые формулы, которые нам потребуются для последующего изложения материала.

Найдем вероятность  $P_n$  того, что при  $n$  независимых испытаниях, исходом которых может быть одно из двух противоположных событий  $A$  или  $\bar{A}$ , событие  $A$  произойдет ровно  $k$  раз.

Вероятность события  $A$  обозначим  $P(A) = P$ , тогда  $P(\bar{A}) = 1 - P$ . Вероятность конкретной комбинации исходов  $n$  независимых испытаний, когда при определенных  $k$  испытаниях появилось событие  $A$ , а при остальных  $n - k$  испытаниях — событие  $\bar{A}$ , в соответствии с теоремой умножения вероятностей равна  $P^k(1 - P)^{n - k}$ , независимо от того, в каком порядке чередовались в данной комбинации исходы событий  $A$  и  $\bar{A}$ . Интересующее нас сложное событие представляет сумму всех исходов  $n$ -кратных испытаний,

дающих  $k$  появлений события  $A$ . Поскольку исходы с разным порядком чередования событий  $A$  и  $\bar{A}$  несовместны, а число исходов, дающих в разной последовательности  $k$  появлений события  $A$  в  $n$  испытаниях, равно числу сочетаний из  $n$  элементов по  $k$ , то в соответствии с теоремой сложения вероятностей

$$P_{n,k} = C_n^k P^k (1 - P)^{n - k}, \quad (2.7)$$

где

$$C_n^k = \frac{n!}{k! (n - k)!}$$

— число сочетаний из  $n$  элементов по  $k$ .

Формула (2.7) носит название биномиального закона.

Рассмотрим теперь две группы зависимых событий  $A_1, A_2, \dots, A_n$  и  $B_1, B_2, \dots, B_n$  каждая из которых составляет полную группу несовместных событий. В этом случае полная вероятность события, т.е. вероятность того, что событие  $A_i$  вообще произойдет совместно с любым из событий  $B_i$  равна

$$P(A_i) = \sum_{j=1}^m P(A_i B_j) = \sum_{j=1}^m P(B_j) P(A_i | B_j). \quad (2.8)$$

На основании формул (2.5) и (2.8) получим

$$P(B_j | A_i) = \frac{P(B_j) P(A_i | B_j)}{P(A_i)} = \frac{P(B_j) P(A_i | B_j)}{\sum_{j=1}^m P(B_j) P(A_i | B_j)}. \quad (2.9)$$

Формула (2.9) носит название формулы Байеса и позволяет определить условные вероятности  $P(B_j | A_i)$ , характеризующие правдоподобность той или иной гипотезы об исходе непосредственно не наблюдаемого явления  $B$  при известном исходе явления  $A$ , если только известны значения вероятностей  $P(B_j)$  исходов  $B_j$  (в сочетании с любым из исходов  $A_i$ ) и значения условных вероятностей  $P(A_i | B_j)$ .

## 2.2. Случайные величины и их вероятностные характеристики

Если случайная величина  $\xi$  может принимать конечное число дискретных значений  $x_i$ , исчерпывающей вероятностной характеристикой ее служит распределение вероятностей этих значений  $P_i$ . (По аналогии со случайными функциями случайные величины обозначим буквами

греческого алфавита, а их конкретные реализации — буквами латинского алфавита).

Если случайная величина  $\xi$  непрерывна и может принимать любое значение на интервале  $[x_{\min}, x_{\max}]$ , то ее статистической характеристикой может служить так называемый интегральный закон распределения  $F(x) = P(\xi < x)$ , определяющий вероятность того, что случайная величина  $\xi$  не превзойдет значение  $x$ . Из определения интегрального закона распределения вытекает следующее очевидное соотношение:

$$P(x_1 \leq \xi < x_2) = F(x_2) - F(x_1) \quad (2.10)$$

где  $P(x_1 \leq \xi < x_2)$  — вероятность того, что случайная величина  $\xi$  не выйдет за пределы интервала  $[x_1, x_2]$ .

Очевидны следующие свойства функции  $F(x)$ :  $F(x)$  — монотонная неубывающая функция;

$$F(-\infty) = 0, \quad F(+\infty) = 1.$$

Если функция  $F(x)$  дифференцируемая, то в качестве вероятностной характеристики случайной величины удобно использовать дифференциальный закон распределения или закон распределения плотности вероятности

$$w(x) = dF(x) / dx \quad (2.11)$$

Очевидны следующие соотношения:

$$F(x) = \int_{-\infty}^x w(x) dx; \quad \int_{-\infty}^{\infty} w(x) dx = 1;$$

$$P(x_1 \leq \xi < x_2) = \int_{x_1}^{x_2} w(x) dx.$$

Помимо законов распределения, часто используются числовые характеристики случайных величин, так называемые *моменты распределения*. Моменты, характеризующие распределение случайных величин относительно нуля, называются начальными.

Для непрерывных случайных величин начальный момент  $k$ -го порядка определяется по формуле

$$m_k(\xi) = \int_{-\infty}^{\infty} x^k w(x) dx. \quad (2.12)$$

При этом предполагается, что несобственный интеграл (2.12) абсолютно сходится, т.е.

$$\int_{-\infty}^{\infty} |x|^k w(x) dx$$

имеет конечное значение.

Для дискретной случайной величины  $\xi$ , принимающей значения  $x_1, x_2, \dots, x_n$  с вероятностями  $P_1, P_2, \dots, P_n$ ,

$$m_k(\xi) = \sum_{r=1}^n x_r^k P_r. \quad (2.13)$$

Наиболее важное значение имеют моменты 1-го и 2-го порядков. Начальный момент 1-го порядка дает *математическое ожидание* или *среднее значение* случайной величины  $\xi$ :

$$m_1(\xi) = \int_{-\infty}^{\infty} x w(x) dx. \quad (2.14)$$

Разность  $\Delta_\xi = \xi - m_1(\xi)$  называется отклонением случайной величины. Моменты распределения отклонений случайной величины называются *центральными* и обозначаются  $M_k(\xi)$ . Нетрудно убедиться, что  $M_1(\xi) = 0$ .

Случайные величины с нулевым средним значением называются *центрированными*. Любые случайные величины можно свести к центрированным, если перейти к отклонениям  $\Delta_\xi$ .

Начальный момент 2-го порядка определяет средний квадрат случайной величины  $\xi$ :

$$m_2(\xi) = \int_{-\infty}^{\infty} x^2 w(x) dx. \quad (2.15)$$

Центральный момент 2-го порядка называется *дисперсией* случайной величины:

$$M_2(\xi) = \sigma_\xi^2 = \int_{-\infty}^{\infty} [x - m_1(\xi)]^2 w(x) dx. \quad (2.16)$$

Дисперсия характеризует разброс случайной величины относительно ее среднего значения.

Величина  $\sigma_\xi = \sqrt{M_2(\xi)}$  называется *среднеквадратическим* или *стандартным отклонением* случайной величины от ее среднего значения.

Отношение  $\Delta_\xi / \sigma_\xi$  называется *нормированным отклонением* случайной величины.

Центральный и начальный момент 2-го порядка случайной величины связаны простым соотношением:

$$M_2(\xi) = m_2(\xi) - m_1^2(\xi). \quad (2.17)$$

Если закон распределения случайной величины симметричен относительно ее среднего значения, т.е.  $w(m_1 + x) = w(m_1 - x)$ , то центральный момент 3-го порядка такой случайной величины равен нулю:

$$\int_{-\infty}^{\infty} (x - m_1)^3 w(x) dx = \int_{-\infty}^{\infty} y^3 w(m_1 + y) dy = 0,$$

так как  $y^3 w(m_1 + y)$  — нечетная функция  $y$ .

Отличие от нуля центрального момента 3-го порядка характеризует асимметрию закона распределения. Безразмерный коэффициент

$$k = M_3 / \sqrt{M_2^3}$$

называют *коэффициентом асимметрии*.

Для совокупности двух случайных величин  $\xi_1$  и  $\xi_2$  исчерпывающей вероятностной характеристикой служит двумерный интегральный закон распределения  $F(x_1, x_2) = P(\xi_1 < x_1, \xi_2 < x_2)$ , определяющий вероятность того, что случайные величины  $\xi_1$  и  $\xi_2$  не превосходят соответственно значений  $x_1$  и  $x_2$ . Если функция  $F(x_1, x_2)$  дифференцируемая, то вероятностной характеристикой двумерной случайной величины может служить двумерный дифференциальный закон распределения или двумерная функция распределения  $w_2(x_1, x_2) = \partial^2 F(x_1, x_2) / \partial x_1 \partial x_2$ . Аналогично могут быть введены многомерный интегральный закон распределения и многомерные функции распределения для совокупностей из любого числа случайных величин.

Можно также ввести числовые характеристики для совокупности двух случайных величин  $\xi_1$  и  $\xi_2$ , имеющих двумерную функцию распределения  $w_2(x_1, x_2)$ . Весьма важной числовой характеристикой совокупности двух случайных величин является смешанный второй центральный момент или ковариация случайных величин  $\xi_1$  и  $\xi_2$ :

$$\begin{aligned} M_2(\xi_1, \xi_2) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x_1 - m_1(\xi_1)] [x_2 - m_1(\xi_2)] w_2(x_1, x_2) \times \\ &\times dx_1 dx_2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 w_2(x_1, x_2) dx_1 dx_2 - m_1(\xi_1) m_1(\xi_2). \end{aligned} \quad (2.18)$$

Если случайные величины  $\xi_1$  и  $\xi_2$  независимы, то  $w_2(x_1, x_2) = w(x_1) w(x_2)$ , и двукратный интеграл в (2.18) распадается на произведение двух однократных интегралов:

$$\int_{-\infty}^{\infty} x_1 w(x_1) dx_1 = m_1(\xi_1), \quad \int_{-\infty}^{\infty} x_2 w(x_2) dx_2 = m_1(\xi_2)$$

и, следовательно,  $M_2(\xi_1, \xi_2) = 0$ . Поэтому ковариация  $M_2(\xi_1, \xi_2)$  может служить некоторой мерой зависимости между двумя случайными величинами. Чаще в качестве такой меры принимают безразмерный коэффициент корреляции

$$\rho_{\xi_1, \xi_2} = M_2(\xi_1, \xi_2) / \sqrt{M_2(\xi_1) M_2(\xi_2)}.$$

Две случайные величины, для которых коэффициент корреляции равен нулю, называются *некоррелированными*. Случайные величины, для которых 1-й начальный момент их произведения

$$m_1(\xi_1 \xi_2) = 0,$$

называют *ортгогональными*. Если средние значения случайных величин равны нулю, то понятия ортогональности и некоррелированности случайных величин совпадают.

Моменты распределения того или иного порядка, являясь важными числовыми характеристиками случайной величины, не являются, однако, их однозначной полной статистической характеристикой: случайные величины, имеющие одинаковые 1-й и 2-й моменты, могут иметь разные законы распределения.

### 2.3. Нормальный закон распределения

Среди случайных величин особое место занимают нормальные случайные величины, подчиняющиеся так называемому нормальному закону распределения:

$$w(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left[ -\frac{(x-a)^2}{2\sigma^2} \right]. \quad (2.19)$$

Определим 1-й начальный момент нормальной случайной величины  $\xi$ :

$$m_1(\xi) = \int_{-\infty}^{\infty} x w(x) dx = \int_{-\infty}^{\infty} \frac{x}{\sigma \sqrt{2\pi}} \exp \left[ -\frac{(x-a)^2}{2\sigma^2} \right] dx.$$

Сделав замену  $(x-a) / \sqrt{2} \sigma = y$ , получим

$$m_1(\xi) = \frac{\sqrt{2}\sigma}{\sqrt{\pi}} \int_{-\infty}^{\infty} y e^{-y^2} dy + \frac{a}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-y^2} dy = a,$$

так как первый из интегралов равен нулю в силу нечетности подынтегральной функции, а



$$\int_{-\infty}^{\infty} e^{-y^2} dy = \sqrt{\pi} \quad [34].$$

Итак,  $a = m_1(\xi)$ .

Определим теперь 2-й центральный момент нормальной случайной величины  $\xi$ :

$$\begin{aligned} M_2(\xi) &= \int_{-\infty}^{\infty} (x-a)^2 \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{x-a}{2\sigma^2}\right] dx = \\ &= \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} y^2 e^{-y^2} dy = \sigma^2, \end{aligned}$$

так как

$$\int_{-\infty}^{\infty} y^2 e^{-y^2} dy = \frac{\sqrt{\pi}}{2} \quad [34].$$

Таким образом, величины  $a$  и  $\sigma^2$ , полностью определяющие нормальный закон распределения, представляют собой соответственно среднее значение и дисперсию случайной величины  $\xi$ , т.е. нормальный закон распределения полностью определяется, если известны первые два момента. Из формулы (2.19) видно, что нормальное распределение симметрично относительно среднего значения случайной величины  $a$ , т.е. 3-й центральный момент и все центральные моменты нечетного порядка равны нулю. Максимум плотности вероятности, соответствующий  $x = a$ , равен

$$w_{\max} = w(a) = 1/\sigma\sqrt{2\pi}.$$

На рис. 2.1 приведен вид нормального закона для различных  $\sigma$  при  $a = 0$ .

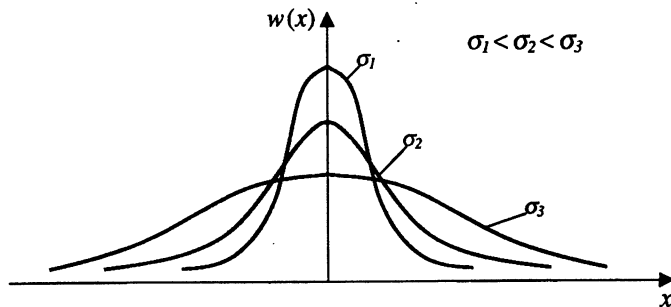


Рис. 2.1

Нормальный закон распределения занимает особое положение в силу того, что большинство реальных случайных величин имеет распределение, близкое к нормальному. Последнее обстоятельство связано с тем, что

на практике случайные величины обычно являются результатом совокупного действия многих независимых случайных факторов и, при некоторых условиях, по мере увеличения числа этих факторов закон распределения асимптотически приближается к нормальному. Условия эти определяются центральной предельной теоремой теории вероятности, которая в упрощенном виде может быть сформулирована следующим образом [1]: если независимые случайные величины  $\xi_1, \xi_2, \dots, \xi_n$  имеют одинаковые распределения с конечной, отличной от нуля дисперсией  $\sigma^2$ , то при  $n \rightarrow \infty$  сумма этих величин стремится к нормальному распределению со средним значением

$$a_{\Sigma} = \sum_{k=1}^n m_1(\xi_k)$$

и дисперсией  $\sigma_{\Sigma}^2 = n\sigma^2$ .

А.М. Ляпунов показал [36], что тенденция к нормализации суммы случайных величин имеет место и при более общих предположениях.

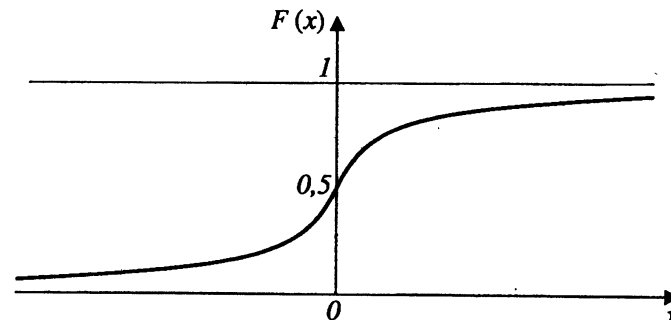


Рис. 2.2

Решение многих практических задач не столь критично к точности аппроксимации закона распределения и уже в случае, когда случайная величина определяется несколькими примерно равноценными независимыми случайными факторами, закон ее распределения можно приближенно аппроксимировать нормальным законом.

Интегральный закон распределения, соответствующий нормальному закону (2.19), имеет вид:

$$F(x) = \int_{-\infty}^x w(y) dy = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x \exp\left[-\frac{(y-a)^2}{2\sigma^2}\right] dy \quad (2.20)$$

Если перейти к нормированным отклонениям  $(x-a)/\sigma$ , то получим

$$F\left(\frac{x-a}{\sigma}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(x-a)/\sigma} e^{-t^2/2} dt. \quad (2.21)$$

Функция

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-t^2/2} dt,$$

представляющая собой вероятность того, что нормированное случайное отклонение не превзойдет величину  $z$ , называется *интегралом вероятности*. Вид этой функции приведен на рис. 2.2.

Поскольку

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-t^2/2} dt = \frac{1}{\sqrt{2\pi}} \int_z^{\infty} e^{-t^2/2} dt = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-t^2/2} dt,$$

то  $F(-z) = 1 - F(z)$  и функцию  $F(z)$  достаточно определить в положительной области.

Вероятность того, что нормальная случайная величина  $\xi$  не выйдет за пределы интервала  $[x_1, x_2]$

$$P(x_1 \leq \xi \leq x_2) = F\left(\frac{x_2 - a}{\sigma}\right) - F\left(\frac{x_1 - a}{\sigma}\right). \quad (2.22)$$

## 2.4. Случайные функции и их вероятностное описание

Случайные функции так же, как и случайные величины, должны определяться вероятностными законами распределения их реализаций. Каждая реализация случайной функции  $\xi(t)$  представляет функцию времени  $x_i(t)$ . Если на спектр случайного сигнала не наложить никаких ограничений, то его реализации даже при конечной их длительности будут определяться бесконечным числом координат, представляющих значения функции  $x_i(t)$  в различные моменты времени. Вероятностное описание такой случайной функции эквивалентно вероятностному описанию бесконечномерной случайной величины и соответственно требует использования бесконечномерных законов распределения.

Задача описания случайной функции значительно упрощается, если ширину ее спектра ограничить сверху величиной  $\mathcal{F}_b$ . В этом случае каждая реализация случайной функции длительностью  $T$  в соответствии с теоремой Котельникова полностью определяется дискретной выборкой, содержащей  $N = 2\mathcal{F}_b T$  отсчетов, следующих с интервалом  $\Delta t = 1/2\mathcal{F}_b$ . Такая случайная функция может быть задана  $N$ -мерным дифференциальным законом распределения, определяющим плотность вероятности

$w_N(x_1, x_2, \dots, x_N)$  ее реализации, характеризуемой совокупностью значений  $x_1, x_2, \dots, x_N$  в дискретных точках отсчета.

Вероятность того, что осуществится реализация, значения которой в дискретных точках отсчета лежат в пределах  $(x_1, x_1 + dx_1), \dots, (x_N, x_N + dx_N)$ , равна  $w_N(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N$ .

Заменяя непрерывные реализации  $x_i(t)$   $N$ -мерной дискретной выборкой, мы тем самым все реализации, имеющие одинаковый спектр до некоторой граничной частоты  $\mathcal{F}_b$  и отличающиеся лишь высокочастотными составляющими спектра за этой границей, заменяем некоторой средней реализацией, соответствующей "усечению" спектров верхней границей  $\mathcal{F}_b$ . Среднеквадратическая ошибка представления всех этих реализаций общей  $N$ -мерной выборкой согласно (1.31) не превосходит  $\sqrt{\Delta E/T}$ , где  $\Delta E$  — энергия, приходящаяся на часть спектра реализаций, лежащую выше границы  $\mathcal{F}_b$ . Выбор размерности закона распределения  $w_N(x_1, x_2, \dots, x_N)$ , характеризующего случайную функцию  $\xi(t)$ , зависит от требований к точности представления ее реализаций. Чем меньше размерность  $N$  закона распределения (чем меньше выбрана граничная частота  $\mathcal{F}_b$ ), тем больше величина  $\Delta E/T$ , характеризующая погрешность представления реализаций случайной функции данной выборкой. И, наоборот, чем больше размерность  $N$  закона распределения, тем выше точность представления реализации случайного сигнала.

## 2.5. Корреляционные характеристики случайных процессов. Стационарные и эргодические случайные процессы

При рассмотрении случайных величин наряду с законами распределения весьма полезными оказались некоторые числовые характеристики (среднее значение, дисперсия и др.), представляющие детерминированные величины. При рассмотрении случайных процессов аналогичную роль играют детерминированные функции, дающие те же числовые характеристики случайных величин, представляющих значения этих процессов в соответствующих временных сечениях. К числу таких характеристик относятся среднее значение, дисперсия, корреляционная функция случайного процесса и др. Среднее значение или первый начальный момент случайного процесса  $\xi(t)$  представляет функцию времени

$$a_{\xi}(t) = m_1\{\xi(t)\} = \int_{-\infty}^{\infty} x w_1(x; t) dx, \quad (2.23)$$

где  $w_1(x; t)$  — одномерное распределение случайной величины, представляющей возможные значения случайного процесса  $\xi(t)$  в фиксированный момент времени  $t$ . Дисперсия случайного процесса  $\xi(t)$  или 2-й центральный момент также представляет функцию времени

$$\sigma_{\xi}^2 = m_1 \{ [\xi(t) - a_{\xi}(t)]^2 \} = \int_{-\infty}^{\infty} [x - a_{\xi}(t)]^2 w_1(x; t) dx. \quad (2.24)$$

Корреляционная функция случайного процесса или смешанный 2-й начальный момент представляет двумерную функцию времени

$$\begin{aligned} B_{\xi}(t_1, t_2) &= m_1 [\xi(t_1) \xi(t_2)] = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 w_2(x_1, x_2; t_1, t_2) dx_1 dx_2, \end{aligned} \quad (2.25)$$

где  $w_2(x_1, x_2; t_1, t_2)$  — двумерное распределение случайных величин  $\xi(t_1)$  и  $\xi(t_2)$ , представляющих возможные значения случайного сигнала  $\xi(t)$  для временных сечений  $t_1$  и  $t_2$ . Часто вместо случайного процесса рассматривают его флуктуации — отклонения от среднего значения

$$\xi_0(t) = \xi(t) - a_{\xi}(t). \quad (2.26)$$

Нетрудно видеть, что  $a_{\xi_0}(t) = 0$ , а  $\sigma_{\xi_0}^2(t) = \sigma_{\xi}^2(t)$ .

Корреляционная функция флуктуации случайного процесса носит название *ковариационной функции* и имеет вид

$$B_{\xi_0}(t_1, t_2) = m_1 \{ [\xi(t_1) - a_{\xi}(t_1)] [\xi(t_2) - a_{\xi}(t_2)] \}. \quad (2.27)$$

Из (2.27) и (2.24) следует, что

$$\sigma_{\xi}^2(t) = B_{\xi_0}(t, t),$$

т.е. дисперсия представляет значение ковариационной функции  $B_{\xi_0}(t_1, t_2)$  при равенстве аргументов  $t_1$  и  $t_2$ .

Случайный процесс  $\xi(t)$  называется *строго стационарным* или *стационарным в "узком смысле"*, если его функция распределения  $w_n(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)$  произвольного порядка  $n$  не меняется при любом сдвиге всей группы точек  $t_1, t_2, \dots, t_n$ , т.е. при соблюдении для любых  $n$  и  $\tau$  условия:

$$w_n(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = w_n(x_1, x_2, \dots, x_n; t_1 + \tau, t_2 + \tau, \dots, t_n + \tau) \quad (2.28)$$

Из (2.28) следует, что двумерная функция распределения стационарного процесса зависит только от интервала  $\tau = t_2 - t_1$  между точками отсчета:

$$w_2(x_1, x_2; t_1, t_2) = w_2(x_1, x_2; \tau), \quad (2.29)$$

а одномерная функция распределения  $w_1(x)$  не зависит от временного сечения.

Соответственно для стационарного случайного процесса от начала отсчета времени не зависят и введенные статистические характеристики

$$a_{\xi}(t) = a_{\xi} = \text{const}; \quad (2.30)$$

$$\sigma_{\xi}^2(t) = \sigma_{\xi}^2 = \text{const}; \quad (2.31)$$

$$B_{\xi}(t, t + \tau) = B_{\xi}(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 w_2(x_1, x_2; \tau) dx_1 dx_2 \quad (2.32)$$

Следует заметить, что условие (2.29) и вытекающие из него условия (2.30) — (2.32) являются необходимыми, но недостаточными условиями строгой стационарности случайного процесса, ибо независимость от начала отсчета времени двумерной функции распределения еще не гарантирует соблюдение этого же условия и для всех распределений более высокого порядка. Однако для большого числа практических задач оказывается достаточным знание свойств процессов, определяемых первыми двумя моментами. Раздел теории, посвященный изучению этих свойств, называется *корреляционной теорией*. В рамках корреляционной теории процессы, удовлетворяющие условию (2.29), ничем не отличаются по своим свойствам от строго стационарных. Случайные процессы, удовлетворяющие этому условию, получили название *стационарных в широком смысле* (или в смысле Хинчина).

В данной книге в основном мы будем рассматривать именно такие процессы и будем называть их для краткости просто *стационарными*.

Корреляционная и ковариационная функции стационарного случайного процесса, описываемого действительной функцией времени, обладают следующими свойствами\*:

1) являются четными функциями аргумента  $\tau$

$$B_{\xi}(\tau) = B_{\xi}(-\tau); \quad B_{\xi_0}(\tau) = B_{\xi_0}(-\tau);$$

\* Случайные процессы, описываемые комплексными функциями времени, здесь не рассматриваются.

2) при  $\tau = 0$  имеют максимальное значение, равное соответственно  $\sigma_\xi^2 + a_\xi^2$  и  $\sigma_\xi^2$ ;

3) при  $\tau \rightarrow \infty$   $B_\xi \rightarrow a_\xi^2$ , а  $B_{\xi_0} \rightarrow 0$ .

Нормированной корреляционной функцией называется функция

$$\rho_\xi = \frac{B_\xi(\tau) - a_\xi^2}{\sigma_\xi^2} = \frac{B_{\xi_0}(\tau)}{B_{\xi_0}(0)}. \quad (2.33)$$

Очевидно, что  $\rho_\xi(0) = 1$  и  $\rho_\xi(\infty) = 0$ .

Из последнего условия следует, что для стационарного случайного процесса, как правило, можно выбрать такое значение  $\tau_0$ , что при  $\tau > \tau_0$  корреляцией между отсчетами процесса можно пренебречь.

Поскольку максимальное значение  $|\rho_\xi(\tau)|$  равно 1, то в качестве меры скорости спада  $|\rho_\xi(\tau)|$ , т.е. меры интервала  $\tau_0$ , при значительном превышении которого можно не считаться с корреляцией между отсчетами, можно использовать площадь под кривой  $|\rho_\xi(\tau)|$ , приняв

$$\tau_0 = \frac{1}{2} \int_{-\infty}^{\infty} |\rho_\xi(\tau)| d\tau. \quad (2.34)$$

(Практически корреляцией можно пренебречь уже при  $\tau \geq (3-4)\tau_0$ ). Величина  $\tau_0$  определяемая формулой (2.34), называется *интервалом корреляции*. Геометрически интервал корреляции может интерпретироваться как половина ширины прямоугольника единичной высоты, площадь которого равновелика площади, ограничиваемой кривой  $|\rho_\xi(\tau)|$  (рис. 2.3).

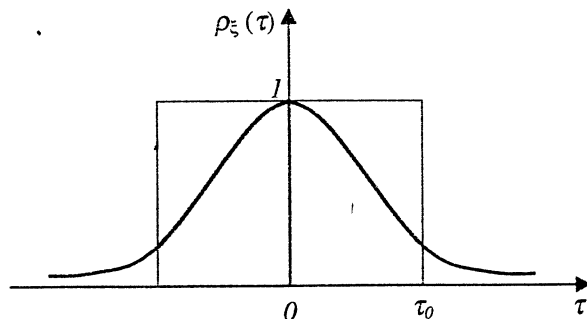


Рис. 2.3

До сих пор мы рассматривали статистические характеристики, получаемые осреднением по множеству реализаций. Рассмотрим теперь некоторые числовые характеристики, получаемые осреднением тех же величин по времени для реализаций большой длительности. Осреднение по времени будем обозначать скобками вида  $\langle \dots \rangle$ . Временное среднее  $r$ -й реализации

$$\langle x_r(t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x_r(t) dt \quad (2.35)$$

можно трактовать как постоянную составляющую реализации.

Средний по времени квадрат реализации

$$\langle x_r^2(t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x_r^2(t) dt \quad (2.36)$$

можно трактовать как среднюю мощность реализации (представляя реализацию как изменение тока или напряжения на нагрузке 1 Ом). Можно ввести функцию временной корреляции

$$\langle x_r(t) x_r(t + \tau) \rangle = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x_r(t) x_r(t + \tau) dt. \quad (2.37)$$

В общем случае для произвольного случайного процесса эти числовые характеристики имеют для разных реализаций различное значение и поэтому не имеют смысла как статистические характеристики процесса. Исключение составляют эргодические случайные процессы.

Случайный процесс называется *эргодическим*, если любая его статистическая характеристика, полученная усреднением по множеству возможных реализаций, с вероятностью, сколь угодно близкой к единице, равна соответствующему временному среднему, полученному на достаточно большом интервале времени из одной единственной реализации случайного процесса. Эргодичность случайного процесса предполагает его стационарность, но не наоборот: стационарный процесс может не быть эргодическим. Например, случайный процесс вида

$$\xi(t) = \xi_0(t) + \eta. \quad (2.38)$$

где  $\xi_0(t)$  — эргодический случайный процесс;  $\eta$  — случайная величина, меняющаяся от реализации к реализации, будет стационарным, но не эргодическим (его числовые характеристики, полученные осреднением по времени, меняются от реализации к реализации из-за изменения  $\eta$ ).

Строгая эргодичность означает, что множество реализаций данного случайного процесса на некотором конечном отрезке времени характеризуется теми же многомерными законами распределения любого порядка, что и множество отрезков этого процесса той же длительности, "нарезанных" из одной бесконечно длинной реализации. Наряду с понятием строгой эргодичности по отношению ко всем возможным вероятностным характеристикам случайного процесса могут использоваться частные определения эргодичности, относящиеся к отдельным характеристикам (эргодичность по флуктуациям случайного процесса, дисперсии и т. д.).

В частности, процесс (2.38), будучи не эргодическим в строгом смысле, эргодичен по отношению к флуктуациям.

Если случайный процесс эргодический, то любая его достаточно длительная реализация представляет свойства всей совокупности. Начальный момент  $k$ -го порядка для такого процесса и его корреляционная функция удовлетворяют условиям:

$$m_k [\xi(t)] = m_1 [\xi^k(t)] = \langle \xi_r^k(t) \rangle; \quad (2.39)$$

$$B_\xi(\tau) = m_1 [\xi(t) \xi(t+\tau)] = \langle \xi_r(t) \xi_r(t+\tau) \rangle$$

для любой  $r$ -й реализации процесса.

Эргодичность процесса создает большие удобства для опытного определения статистических характеристик, так как позволяет получить их из одной достаточно длительной реализации случайного процесса.

## 2.6. Многомерный нормальный закон распределения.

### Нормальные случайные процессы

Рассмотрим многомерный нормальный закон распределения, имеющий вид [1, 16]:

$$w_n(x_1, x_2, \dots, x_n) = \frac{1}{\sigma_1 \sigma_2 \dots \sigma_n \sqrt{(2\pi)^n D}} \times \\ \times \exp \left[ -\frac{1}{2D} \sum_{i=1}^n \sum_{k=1}^n A_{ik} \frac{(x_i - a_i)(x_k - a_k)}{\sigma_i \sigma_k} \right], \quad (2.40)$$

где

$$D = \begin{vmatrix} 1 & \rho_{12} & \dots & \rho_{1n} \\ \rho_{21} & 1 & \dots & \rho_{2n} \\ \dots & \dots & \dots & \dots \\ \rho_{n1} & \rho_{n2} & \dots & 1 \end{vmatrix}$$

— определитель<sup>\*)</sup>;  $A_{ik}$  — алгебраическое дополнение элемента  $\rho_{ik}$  определителя;  $\rho_{ik} = \rho_{ki}$  — некоторые числовые величины, удовлетворяющие условию  $|\rho_{ik}| < 1$  при  $i \neq k$ . В частности, при  $n = 2$ , обозначая  $\rho_{12} = \rho_{21} = \rho$ , получаем  $D = 1 - \rho^2$ ,  $A_{12} = A_{21} = \rho$ ,  $A_{11} = A_{22} = 1$  и двумерный нормальный закон распределения случайных величин  $\xi_1, \xi_2$  принимает вид:

$$w_2(x_1, x_2) = \frac{1}{\sigma_1 \sigma_2 2\pi \sqrt{1 - \rho^2}} \exp \left\{ -\frac{1}{2(1 - \rho^2)} \left[ \frac{(x_1 - a_1)^2}{\sigma_1^2} - \right. \right. \\ \left. \left. - 2\rho \frac{(x_1 - a_1)(x_2 - a_2)}{\sigma_1 \sigma_2} + \frac{(x_2 - a_2)^2}{\sigma_2^2} \right] \right\}. \quad (2.41)$$

Для того, чтобы раскрыть смысл величин  $a_1, a_2, \sigma_1, \sigma_2$  входящих в (2.41), достаточно определить одномерные законы распределения случайных величин  $\xi_1$  и  $\xi_2$ :

$$w_1(x_1) = \int_{-\infty}^{\infty} w_2(x_1, x_2) dx_2,$$

$$w_1(x_2) = \int_{-\infty}^{\infty} w_2(x_1, x_2) dx_1.$$

Для функции  $w_2(x_1, x_2)$  вида (2.41) интегралы в правых частях этих выражений сводятся к табличным и одномерные законы распределения приводятся к виду [1]:

$$w_1(x_1) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp \left[ -\frac{(x_1 - a_1)^2}{2\sigma_1^2} \right],$$

$$w_1(x_2) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp \left[ -\frac{(x_2 - a_2)^2}{2\sigma_2^2} \right].$$

Таким образом, одномерные законы распределения случайных величин  $\xi_1$  и  $\xi_2$  получаются нормальными вида (2.19), а величины  $\sigma_1^2, a_1$  и  $\sigma_2^2, a_2$  представляют их дисперсии и средние значения.

Для определения величины  $\rho$ , входящей в (2.41), нужно найти ковариацию случайных величин  $\xi_1$  и  $\xi_2$ :

$$M_2(\xi_1, \xi_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 - a_1)(x_2 - a_2) w_2(x_1, x_2) dx_1 dx_2.$$

Для функции  $w_2(x_1, x_2)$  вида (2.41) этот двойной интеграл сводится к произведению табличных однократных интегралов и выражение для ковариации случайных величин  $\xi_1$  и  $\xi_2$  приводится к виду [1]:

$$M_2(\xi_1, \xi_2) = \rho \sigma_1 \sigma_2.$$

Откуда

<sup>\*)</sup> Краткие сведения об определителях см. в приложении.

$$\rho = M_2(\xi_1, \xi_2) / \sigma_1 \sigma_2. \quad (2.42)$$

Таким образом, величина  $\rho$  представляет коэффициент корреляции между случайными величинами, определяемыми нормальным двумерным законом распределения (2.41).

Аналогично можно показать, что и в  $n$ -мерном нормальном законе распределения (2.40) величины  $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$  представляют дисперсии  $a_1, a_2, \dots, a_n$  — средние значения,  $\rho_{ij}$  — коэффициенты корреляции случайных величин  $\xi_1, \xi_2, \dots, \xi_n$  определяемых законом распределения. В частном случае, когда случайные величины не коррелированы и  $\rho_{ij} = 0, i \neq j$ , нормальный многомерный закон распределения принимает вид

$$w_n(x_1, x_2, \dots, x_n) = \prod_{i=1}^n \frac{1}{\sigma_i \sqrt{2\pi}} \exp \left[ -\frac{(x_i - a_i)^2}{2\sigma_i^2} \right], \quad (2.43)$$

т.е. представляет произведение одномерных нормальных законов распределения.

Таким образом, если случайные величины подчиняются многомерному нормальному распределению, то их некоррелированность является достаточным условием их независимости. Для произвольного закона распределения это несправедливо.

Как уже отмечалось, реализации случайного процесса с требуемой точностью могут определяться дискретными выборками отсчетов, а сам случайный процесс может определяться многомерным законом распределения случайных величин, составляющих выборку. Случайный процесс  $\xi(t)$  называется *нормальным (гауссовым)*, если законы распределения любого  $n$ -го порядка для совокупности случайных величин  $\xi(t_k), k=1, 2, \dots, n$ , нормальные, т.е. определяются формулой (2.40). Поскольку величины  $a_i, \sigma_i$  и  $\rho_{ij}$ , входящие в формулу (2.40), для любой выборки дискретных отсчетов представляют значения функций  $a_\xi(t_i)$  и  $B_\xi(t_i, t_j)$  в соответствующих временных сечениях, то нормальный случайный процесс полностью определяется этими двумя характеристиками. Поэтому для нормальных процессов условия стационарности в широком смысле достаточны и для определения строгой стационарности и корреляционная теория имеет неограниченную область применения. Для нормального стационарного процесса

$$a(t) = a_\xi = \text{const}; \quad B_\xi(t_1, t_2) = B_\xi(\tau);$$

$$\sigma_\xi^2 = B_{\xi_0}(0) = B_\xi(0) - a_\xi^2 = \text{const};$$

$$\rho_\xi(\tau) = B_{\xi_0}(\tau) / \sigma_\xi^2,$$

и при использовании выборки отсчетов с интервалом дискретности  $\Delta t = T/n$   $n$ -мерный закон распределения, определяющий этот процесс, имеет вид:

$$w_n(x_1, x_2, \dots, x_n) = \frac{1}{\sigma_\xi^n (2\pi)^{n/2} \sqrt{D}} \times \\ \times \exp \left[ -\frac{1}{2D} \sum_{i=1}^n \sum_{j=1}^n A_{ij} \frac{(x_i - a_\xi)(x_j - a_\xi)}{\sigma_\xi^2} \right], \quad (2.44)$$

где входящие в определитель  $D$  величины равны

$$\rho_{ij} = B_{\xi_0}[(i-j)\Delta t] / \sigma_\xi^2. \quad (2.45)$$

## 2.7. Спектральное представление стационарных случайных процессов

В качестве спектральной характеристики стационарного случайного процесса будем использовать осредненное по множеству реализаций значение спектральной плотности мощности.

Рассмотрим одну из реализаций  $x_k(t)$  стационарного случайного процесса  $\xi(t)$ . Заменим ее усеченной реализацией  $x_{kT}(t)$ , равной нулю вне интервала  $|t| \leq T/2$ . В соответствии с (1.10), (1.11) усеченная реализация  $x_{kT}(t)$  может быть разложена в ряд Фурье с интервалом между гармониками  $\Delta f = 1/T$  и комплексной амплитудой гармоники  $f_i$

$$\dot{C}_{kT}(f_i) = \frac{1}{T} \int_{-T/2}^{T/2} x_{kT}(t) e^{-j2\pi f_i t} dt. \quad (2.46)$$

Средняя мощность, приходящаяся на единицу полосы спектра реализации  $x_{kT}(t)$ , при этом равна

$$G_{kT}(f) = \frac{1}{\Delta f} |\dot{C}_{kT}(f)|^2 = T \dot{C}_{kT}(f) \dot{C}_{kT}(-f) = \\ = T \frac{1}{T^2} \int_{-T/2}^{T/2} x_{kT}(t_1) e^{-j2\pi f t_1} dt_1 \times \int_{-T/2}^{T/2} x_{kT}(t_2) e^{-j2\pi f t_2} dt_2 = \\ = \frac{1}{T} \int_{-T/2}^{T/2} \int_{-T/2}^{T/2} x_{kT}(t_1) x_{kT}(t_2) e^{-j2\pi f(t_1 - t_2)} dt_1 dt_2. \quad (2.47)$$

Осредним  $G_{\xi T}(f)$  по множеству реализаций, положив процесс стационарным\* (операцию осреднения обозначим чертой сверху):

$$\overline{G_T}(f) = \frac{1}{T} \int_{-T/2}^{T/2} \int_{-T/2}^{T/2} B_{\xi T}(t_1 - t_2) e^{-j2\pi f(t_1 - t_2)} dt_1 dt_2.$$

Введя вместо  $t_1$  новую переменную  $\tau = t_1 - t_2$ , получим

$$\overline{G_T}(f) = \frac{1}{T} \int_{-T}^T B_{\xi T}(\tau) e^{-j2\pi f\tau} \int_{t_{2\min}(\tau)}^{t_{2\max}(\tau)} dt_2 d\tau. \quad (2.48)$$

В этом выражении внешний интеграл (по  $\tau$ ) имеет пределы интегрирования от  $-T$  до  $T$ , так как

$$\tau_{\min} = t_{1\min} - t_{2\max} = -T/2 - T/2 = -T,$$

$$\tau_{\max} = t_{1\max} - t_{2\min} = T/2 + T/2 = T.$$

Функция  $B_{\xi T}(t_1 - t_2) = \xi_T(t_1) \xi_T(t_2)$  отлична от нуля лишь при условии, что оба значения аргумента  $t$  принадлежат области  $[-T/2, T/2]$  ненулевых значений функции  $\xi_T(t)$ . Поэтому пределы интегрирования для внутреннего интеграла определяются как крайние значения переменной интегрирования  $t_2$  при фиксированном  $\tau$ , удовлетворяющие этому условию, а именно:

для  $\tau > 0$ , т.е. при  $t_1 > t_2$ ,

$$t_{2\min}(\tau) = t_{\min} = -T/2; \quad t_{2\max}(\tau) = t_{\max} - |\tau| = T/2 - |\tau|;$$

для  $\tau < 0$ , т.е. при  $t_1 < t_2$ ,

$$t_{2\min}(\tau) = t_{\min} + |\tau| = -T/2 + |\tau|; \quad t_{2\max}(\tau) = t_{\max} = T/2.$$

В обоих случаях для разности  $[t_{2\max}(\tau) - t_{2\min}(\tau)]$  получается одинаковое значение

$$[t_{2\max}(\tau) - t_{2\min}(\tau)] = T - |\tau|.$$

Поэтому (2.48) принимает вид

$$\overline{G_T}(f) = \int_{-T}^T B_{\xi T}(\tau) \left(1 - \frac{|\tau|}{T}\right) e^{-j2\pi f\tau} d\tau.$$

Переходя к пределу при  $T \rightarrow \infty$ , получаем

$$\begin{aligned} G_{\xi} &= \lim_{T \rightarrow \infty} \overline{G_T}(f) = \int_{-\infty}^{\infty} B_{\xi}(\tau) e^{-j2\pi f\tau} d\tau = \\ &= \int_{-\infty}^{\infty} B_{\xi}(\tau) \cos 2\pi f\tau d\tau - j \int_{-\infty}^{\infty} B_{\xi}(\tau) \sin 2\pi f\tau d\tau. \end{aligned} \quad (2.49)$$

В силу четности  $B_{\xi}(f)$  второй интеграл равен нулю и

$$G_{\xi}(f) = \int_{-\infty}^{\infty} B_{\xi}(\tau) \cos 2\pi f\tau d\tau = 2 \int_0^{\infty} B_{\xi}(\tau) \cos 2\pi f\tau d\tau, \quad (2.50)$$

т.е.  $G_{\xi}(f)$  — также четная функция.

Функцию частоты  $G_{\xi}(f)$ , представляющую осредненную по множеству реализаций спектральную плотность мощности стационарного процесса, называют *энергетическим спектром*. Это название, укоренившееся в литературе по радиотехнике, которым будем пользоваться и мы, нельзя признать удачным, ибо оно не соответствует фактической размерности величины  $G_{\xi}(f)$ . Если равенство Парсваля (1.18), полученное для детерминированных сигналов с конечной энергией, действительно содержит энергетический спектр, т.е. спектральную плотность энергии сигнала, то для стационарных случайных сигналов мы принципиально не можем воспользоваться аналогичной характеристикой. Стационарный случайный сигнал, строго говоря, имеет неограниченную длительность (иначе нарушались бы условия стационарности). Поэтому при конечной средней мощности он обладает бесконечной средней энергией, а в случае ограниченной ширины спектра — и бесконечной спектральной плотностью энергии. Поэтому для характеристики стационарных случайных сигналов используется спектральная плотность мощности. Заметим, что сигналы с ограниченной энергией, для которых получено равенство Парсваля (1.18), можно характеризовать только спектральной плотностью энергии, ибо спектральная плотность мощности для них обращается в нуль. Неудачный термин "энергетический спектр" в данном случае привился, по-видимому, потому что размерность спектральной плотности мощности совпадает с размерностью энергии (но не спектральной плотности энергии!):

$$G_T(f) = \mathcal{P}/\Delta f = \mathcal{P}T.$$

Из (2.50) следует, что энергетический спектр  $G_{\xi}(f)$  и корреляционная функция  $B_{\xi}(\tau)$  стационарного случайного процесса  $\xi(t)$  связаны преобразованиями Фурье, и поэтому можно записать обратное преобразование:

$$B_{\xi}(f) = \int_{-\infty}^{\infty} G_{\xi}(f) \cos 2\pi f\tau df = \int_0^{\infty} 2G_{\xi}(f) \cos 2\pi f\tau df. \quad (2.51)$$

Таким образом, энергетический спектр стационарного случайного процесса однозначно определяет его корреляционную функцию, а если этот процесс нормальный с известным средним значением, то он определяет и все его статистические свойства. Полученный результат, показывающий, что корреляционная функция и энергетический спектр стационарного случайного сигнала связаны между собой так же, как детерминированный сигнал и его спектр, носит название *теоремы Хинчина—Винера*.

Величина  $G_{\xi}(f)$ , определяемая по формуле (2.50), даст спектральную плотность средней мощности, отвечающую случаю использования двустороннего комплексного спектра. Для одностороннего спектра, отвечающего гармоническим базисным функциям, спектральная плотность

$$\mathcal{G}_{\xi}(f) = 2G_{\xi}(f). \quad (2.52)$$

\* Кроме того, по-прежнему предполагается, что процесс является действительной функцией времени.

При  $\tau = 0$  из (2.51) получаем очевидное соотношение, связывающее среднюю мощность и ее спектральную плотность:

$$B_{\xi}(0) = \int_{-\infty}^{\infty} G_{\xi}(f) df = \int_0^{\infty} \mathcal{G}_{\xi}(f) df. \quad (2.53)$$

Величину площади под кривой одностороннего энергетического спектра, отнесенную к спектральной плотности на некоторой характерной частоте  $f_0$  (обычно соответствующей максимуму спектральной плотности), называют шириной полосы энергетического спектра:

$$\Delta f = \frac{1}{\mathcal{G}(f_0)} \int_0^{\infty} \mathcal{G}_{\xi}(f) df. \quad (2.54)$$

Она равна основанию прямоугольника, имеющего высоту  $\mathcal{G}(f_0)$ , равновеликого по площади средней мощности сигнала  $B_{\xi}(0)$  (рис. 2.4). На основании (2.50) спектральную плотность средней мощности при  $f = 0$  можем записать в виде:

$$\mathcal{G}_{\xi}(0) = 2G_{\xi}(0) = 2 \int_{-\infty}^{\infty} B_{\xi}(\tau) d\tau. \quad (2.55)$$

Если случайный процесс эргодический, то осреднение по множеству реализаций можно заменить осреднением по времени:

$$B_{\xi}(0) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} [\xi_k(t)]^2 dt = \int_{-\infty}^{\infty} G_{\xi}(f) df. \quad (2.56)$$

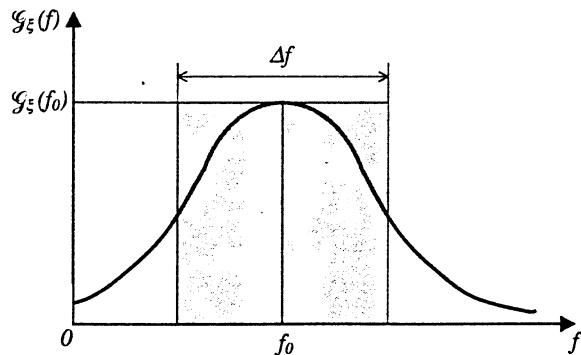


Рис. 2.4

Преобразования Фурье (2.50), (2.51) применимы только для стационарных процессов с нулевым средним значением, так как в противном случае  $B_{\xi}(\tau) \rightarrow a^2_{\xi}$  при  $\tau \rightarrow \infty$  и не выполняется условие абсолютной интегрируемости функции  $B_{\xi}(\tau)$ :

$$\int_{-\infty}^{\infty} |B_{\xi}(\tau)| d\tau \leq M, \quad (2.57)$$

необходимое для применения преобразования Фурье.

Корреляционная функция  $B_{\xi}(\tau)$  и энергетический спектр  $G(f)$ , как всякая пара функций, связанных преобразованием Фурье, подчиняются следующей закономерности: чем шире спектр, тем "уже" корреляционная функция (меньше время корреляции), и наоборот.

## 2.8. Случайный процесс типа белого шума

Рассмотрим центрированный стационарный случайный процесс с равномерным энергетическим спектром в ограниченной полосе:

$$G(f) = \begin{cases} N_0/2 & \text{при } |f| \leq \mathcal{F}, \\ 0 & \text{при } |f| > \mathcal{F}. \end{cases} \quad (2.58)$$

В соответствии с (2.51) корреляционная функция такого случайного процесса имеет вид:

$$B_{\xi}(\tau) = \frac{N_0}{2} \int_{-\mathcal{F}}^{\mathcal{F}} \cos 2\pi f \tau df = \frac{N_0 \sin 2\pi \mathcal{F} \tau}{2\pi \tau}. \quad (2.59)$$

Дисперсия (средняя мощность) этого процесса

$$\sigma_{\xi}^2 = B_{\xi}(0) = N_0 \mathcal{F}. \quad (2.60)$$

При фиксированной спектральной плотности  $N_0$  дисперсия (2.60) неограниченно увеличивается с расширением спектра. Нормированная корреляционная функция процесса, задаваемого энергетическим спектром (2.58), равна

$$\rho_{\xi}(\tau) = B_{\xi}(\tau) / \sigma_{\xi}^2 = \sin 2\pi \mathcal{F} \tau / 2\pi \mathcal{F} \tau. \quad (2.61)$$

Согласно теореме Котельникова реализации случайного процесса с ограниченным спектром полностью определяются выборкой случайных величин, представляющих его дискретные отсчеты через интервалы  $\Delta t = 1/2\mathcal{F}$ . Как следует из (2.61), при  $\tau = k\Delta t = k/2\mathcal{F}$   $\rho_{\xi}(\tau) = 0$ , т.е. отсчеты через равномерные интервалы  $\Delta t$ , принадлежащие выборке, полностью определяют стационарный случайный процесс с равномерным энергетическим спектром, будут некоррелированы. Если стационарный процесс, обладающий энергетическим спектром вида (2.58), является к тому же нормальным, то из некоррелированности отсчетов, составляющих выборку, следует их независимость (см. § 2.6). Такой процесс полностью определяется одномерным нормальным законом распределения его значений в произвольном временном сечении.

Эти свойства сохраняются при любой ширине равномерного энергетического спектра случайного процесса. В пределе при  $\mathcal{F} \rightarrow \infty$  и  $\Delta t \rightarrow 0$  случайный процесс с равномерным энергетическим спектром характеризуется отсутствием корреляции между сколь угодно близкими его отсчетами. Такой случайный процесс носит название "белого" шума. Корреляционная функция белого шума удовлетворяет условию:



$$B_{\xi}(\tau) = \begin{cases} \infty & \text{при } \tau = 0, \\ 0 & \text{при } \tau \neq 0, \end{cases}$$

а его дисперсия (средняя мощность)  $\sigma_{\xi}^2 = B_{\xi}(0) = \infty$ .

Корреляционная функция белого шума математически может быть определена с помощью  $\delta$ -функции, обладающей следующими свойствами:

$$\delta(\tau) = \begin{cases} \infty & \text{при } \tau = 0, \\ 0 & \text{при } \tau \neq 0; \end{cases} \quad (2.62)$$

$$\int_{-\infty}^{\infty} \delta(\tau) d\tau = 1. \quad (2.63)$$

Простейшей моделью  $\delta$ -функции может служить функция получающаяся в результате предельного сжатия прямоугольного импульса единичной площади шириной  $\alpha$  с амплитудой  $1/\alpha$ , симметричного относительно оси ординат (рис. 2.5). Если семейство этих импульсов представить однопараметрической функцией вида  $\Delta(\tau, \alpha)$ , то можно записать, что

$$\delta(\tau) = \lim_{\alpha \rightarrow 0} \Delta(\tau, \alpha). \quad (2.64)$$

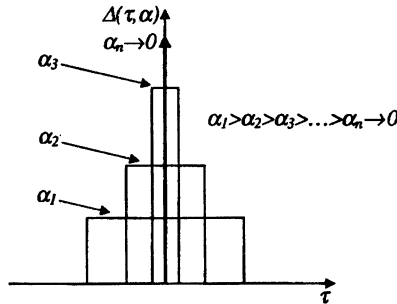


Рис. 2.5

Можно назвать целый ряд других семейств однопараметрических функций, удовлетворяющих условию (2.64). В частности, однопараметрическое семейство одномерных нормальных законов распределения вида

$$w(x, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

удовлетворяет условию (2.64):

$$\lim_{\sigma \rightarrow 0} w(x, \sigma) = \delta(x).$$

Таким образом,  $\delta$ -функция представляет собой обобщенную функцию, получающуюся предельным переходом вида (2.64) от семейств однопараметрических функций  $\Delta(\tau, \alpha)$  различной формы, обладающих общим свойством сжатия при  $\alpha \rightarrow 0$  в бесконечно узкий пик единичной площади.

В случае, когда  $\delta$ -функция стоит под знаком определенного интеграла в качестве множителя, она обладает своеобразным "фильтрующим" свойством: в результате интегриро-

вания "выделяется" значение стоящей при ней в качестве множителя функции, отвечающее величине переменной интегрирования, обращающей в нуль аргумент  $\delta$ -функции. Действительно, если функция  $f(x)$  непрерывна в бесконечно малой  $\epsilon$ -окрестности точки  $x_1$ , то учитывая, что  $\delta(x - x_1) = 0$  при  $|x - x_1| > \epsilon$ , получаем

$$\begin{aligned} \int_{-\infty}^{\infty} f(x) \delta(x - x_1) dx &= \int_{x_1 - \epsilon}^{x_1 + \epsilon} f(x) \delta(x - x_1) dx = \\ &= f(x_1) \int_{x_1 - \epsilon}^{x_1 + \epsilon} \delta(x - x_1) dx = f(x_1). \end{aligned} \quad (2.65)$$

Используя  $\delta$ -функцию, корреляционную функцию "белого шума" можно представить в виде

$$B_{\xi}(\tau) = \frac{N_0}{2} \delta(\tau). \quad (2.66)$$

Легко убедиться, что корреляционной функции вида (2.66) соответствует равномерный энергетический спектр со спектральной плотностью  $G(f) = N_0/2$  на всем интервале частот от  $-\infty$  до  $+\infty$ . Используя (2.50) и учитывая "фильтрующее" свойство  $\delta$ -функции получаем

$$G(f) = \int_{-\infty}^{\infty} B_{\xi}(\tau) e^{-j2\pi f\tau} d\tau = \frac{N_0}{2} \int_{-\infty}^{\infty} \delta(\tau) e^{-j2\pi f\tau} d\tau = \frac{N_0}{2}.$$

Белый шум, определяемый как случайный процесс, отвечающий бесконечно широкому равномерному энергетическому спектру, представляет абстракцию, физически не реализуемую. Его нельзя задать математически как случайный процесс: нельзя задать одномерный закон распределения, определяющий нормальный белый шум, поскольку дисперсия его бесконечна. Мы можем лишь характеризовать локальные свойства белого шума в частотном и временном измерениях пространства сигнала: однородность спектральной плотности мощности и некоррелированность сколь угодно близких по времени отсчетов. Однако эта абстракция оправдана тем, что специфические свойства белого шума переносятся и на случайные процессы с ограниченной мощностью (дисперсией), отвечающие ограниченному участку равномерного спектра, выделяемому входным устройством. Хотя при этом время корреляции шума уже конечно, отсчеты многомерной выборки, следующие через интервал  $\Delta t = 1/2\mathcal{F}$  и полностью задающие шум при ограниченной ширине его спектра, как было показано, будут некоррелированными. Таким образом, корреляционные свойства белого шума полностью переносятся на дискретную выборку, определяющую стационарный процесс, обладающий равномерным спектром в полосе, ограничиваемой входным устройством реальных систем. Корреляционная функция и форма энергетического спектра шума на выходе системы, при действии на его входе белого шума, полностью определяются параметрами системы с точностью до масштабного коэффициента  $N_0$  характеризующего уровень входного шума. Это, кстати, делает белый шум весьма удобным эталонным сигналом при определении характеристик линейной системы (импульсной переходной функции, передаточной функции, ширины полосы пропускания).

Для того чтобы шум, действующий на входе системы, мог рассматриваться как "белый", достаточно, чтобы он имел относительно равномерный энергетический спектр в пределах полосы пропускания входного устройства системы; характер энергетического спектра

шума вне этой полосы не имеет значения. Таким образом, благодаря ограниченной полосе пропускания реальных систем белый шум обретает физическую реальность.

## 2.9. Узкополосные случайные сигналы

Стационарный случайный сигнал называется узкополосным, если его односторонний энергетический спектр  $\mathcal{G}_\xi(f)$  сосредоточен в окрестности некоторой частоты  $f_0$  в относительно узкой полосе:  $\Delta f \ll f_0$  (рис. 2.6).

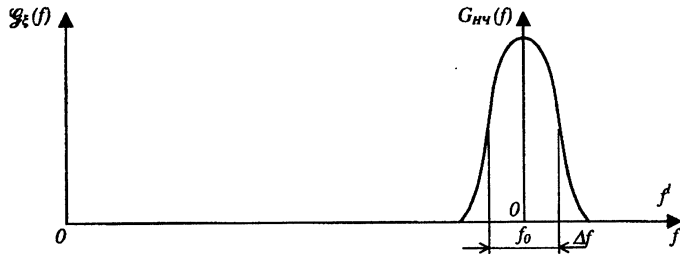


Рис. 2.6

Выразим корреляционную функцию такого случайного сигнала через его односторонний энергетический спектр  $\mathcal{G}_\xi(f)$ :

$$B_\xi(\tau) = \int_0^\infty \mathcal{G}_\xi(f) \cos 2\pi f \tau df.$$

Перейдем от абсолютного значения частоты  $f$  к расстройке  $f' = f - f_0$ . Тогда получим

$$B_\xi(\tau) = \int_{-f_0}^\infty G_{\text{нч}}(f') \cos [2\pi (f_0 + f')\tau] df', \quad (2.67)$$

где  $G_{\text{нч}}(f') = \mathcal{G}(f_0 + f')$  — низкочастотный узкополосный энергетический спектр, соответствующий переносу начала отсчета частоты в точку  $f_0$  (рис. 2.6). Так как расстройка  $f'$  в области  $\Delta f$  ненулевых значений энергетического спектра меняет знак, то при переносе начала отсчета в точку  $f_0$  односторонний энергетический спектр  $\mathcal{G}(f)$  трансформируется в двусторонний спектр  $G_{\text{нч}}(f')$ , сосредоточенный в области  $|f'| \leq \Delta f/2 \ll f_0$ . Поскольку для значений  $f' < -f_0$ , лежащих далеко за границами области ненулевых значений спектра, можно принять  $G_{\text{нч}}(f') = 0$ , то нижний предел интегрирования в (2.67) можно расширить до  $-\infty$  и записать

$$B_\xi(\tau) = \left[ \int_{-\infty}^\infty G_{\text{нч}}(f') \cos (2\pi f' \tau) df' \right] \cos 2\pi f_0 \tau - \left[ \int_{-\infty}^\infty G_{\text{нч}}(f') \sin (2\pi f' \tau) df' \right] \sin 2\pi f_0 \tau. \quad (2.68)$$

Если энергетический спектр симметричен относительно нового начала отсчета частоты, то  $G_{\text{нч}}(f')$  — четная функция  $f'$  и интеграл во втором слагаемом правой части (2.68) обращается в нуль. В этом случае получаем

$$B_\xi(\tau) = \left[ \int_{-\infty}^\infty G_{\text{нч}}(f') \cos (2\pi f' \tau) df' \right] \cos 2\pi f_0 \tau = B_{\text{нч}}(\tau) \cos 2\pi f_0 \tau, \quad (2.69)$$

где  $B_{\text{нч}}(\tau)$  — корреляционная функция огибающей, определяемая низкочастотным энергетическим спектром  $G_{\text{нч}}(f')$ .

Таким образом, корреляционная функция узкополосного сигнала, энергетический спектр которого симметричен относительно частоты  $f_0$  может быть представлена в виде амплитудно-модулированного колебания с частотой  $f_0$ , огибающая которого представляет корреляционную функцию огибающей рассматриваемого узкополосного сигнала.

## Глава 3 СТАТИСТИЧЕСКАЯ ТЕОРИЯ ОБНАРУЖЕНИЯ СИГНАЛОВ

### 3.1. Классификация задач обнаружения сигналов

Сформулируем задачу обнаружения сигналов. Пусть имеется некоторое число различных сигналов  $u_i(t)$ , лишь один из которых может передаваться на интервале наблюдения. Сигналы могут быть либо детерминированными, либо квазидетерминированными, либо случайными с заданными вероятностными характеристиками. На вход приемного устройства поступает реализация случайного сигнала, представляющего смесь переданного полезного сигнала  $u_i(t)$  и случайной помехи  $n(t)$

$$y(t) = u_i(t) + n(t). \quad (3.1)$$

При этом одна и та же реализация суммарного сигнала  $y(t)$  может получиться для различных вариантов переданного сигнала в сочетании с некоторой реализацией случайной помехи

$$n(t) = y(t) - u_i(t). \quad (3.2)$$

Таким образом, после приема сигнала  $y(t)$  могут быть выдвинуты разные гипотезы о том, какой из сигналов  $u_i(t)$  был передан. Задача статистической теории обнаружения сигналов — дать метод выбора наиболее достоверной из этих гипотез.

В зависимости от числа возможных вариантов передаваемых сигналов задачи обнаружения можно подразделить на *бинарные* и *многоальтернативные*.

В бинарных задачах на интервале наблюдения может передаваться один из двух сигналов. Частным случаем бинарной задачи является обнаружение факта передачи или отсутствия единственного сигнала, так как отсутствие сигнала на интервале наблюдения эквивалентно нулевому сигналу. К числу бинарных задач обнаружения относятся задачи обнаружения радиолокационного сигнала, распознавания символов двоичного кода и др.

В многоальтернативных задачах требуется идентифицировать принятый сигнал с одним из  $p$  сигналов с известными характеристиками, где

$p > 2$ . К числу таких задач относится распознавание символов кода с основанием более двух, прием функциональных команд, переданных по командной радиолинии и др.

По характеру полезных сигналов можно выделить следующие классы задач обнаружения:

- 1) обнаружение сигналов с известными параметрами на фоне шума (цифровые системы связи при наличии синхронизации, командные радиолинии др.);
- 2) обнаружение сигналов с неизвестными параметрами на фоне шума (радиолокационное обнаружение, цифровые системы связи без канала синхронизации)
- 3) обнаружение случайных сигналов на фоне шума (сейсмические системы обнаружения, радиоастрономия, радиоразведка).

### 3.2. Общая постановка бинарной задачи обнаружения. Критерий Байеса.

Результат наблюдения на интервале  $T$  реализации  $y(t)$ , представляющей смесь полезного сигнала и шума, можно характеризовать выборкой из  $N$  значений  $(y_1, y_2, \dots, y_N)$ , т.е. представить в виде вектора  $y$  в  $N$ -мерном пространстве. Обозначим через  $H_0$  и  $H_1$  гипотезы, соответствующие двум возможным вариантам передаваемого сигнала  $u_0$  и  $u_1$  (в случае обнаружения единственного сигнала  $H_0$  — гипотеза о его отсутствии, а  $H_1$  — гипотеза о его наличии). В бинарной задаче верна либо гипотеза  $H_0$ , либо гипотеза  $H_1$ . Поэтому при решении задачи обнаружения возможен один из четырех исходов:

- 1) выбрана гипотеза  $H_0$  и верна гипотеза  $H_0$ ;
- 2) выбрана гипотеза  $H_1$ , а верна  $H_0$ ;
- 3) выбрана гипотеза  $H_1$  и верна  $H_1$ ;
- 4) выбрана гипотеза  $H_0$ , а верна  $H_1$ .

Первый и третий исходы соответствуют правильным выборам, второй и четвертый — неверным.

Пусть известны априорные вероятности  $P_0$  и  $P_1$  передачи сигналов  $u_0$  и  $u_1$ . Припишем каждому из четырех возможных исходов решения задачи обнаружения некоторую стоимость, характеризующую те потери (или выигрыш), которые будут при этом исходе. Обозначим эти стоимости соответственно  $C_{00}$ ,  $C_{10}$ ,  $C_{11}$  и  $C_{01}$  (первый индекс обозначает выбранную гипотезу, а второй — правильную). Будем строить наше правило решения таким образом, чтобы средние потери были минимальными (критерий Байеса). Величину средних потерь назовем средним риском и обозначим

буквой  $R$ . Обозначим условные вероятности перечисленных исходов:  $P(H_0|H_0)$ ,  $P(H_1|H_0)$ ,  $P(H_1|H_1)$ , и  $P(H_0|H_1)$ , где первой стоит выбранная гипотеза, а второй — правильная. Тогда средний риск

$$R = C_{00}P_0P(H_0|H_0) + C_{10}P_0P(H_1|H_0) + C_{11}P_1P(H_1|H_1) + C_{01}P_1P(H_0|H_1). \quad (3.3)$$

Примем, что стоимости потерь при ошибочном решении  $C_{10} > 0$  и  $C_{01} > 0$ , тогда как при правильном решении мы имеем выигрыш или отрицательные потери:  $C_{00} \leq 0$  и  $C_{11} \leq 0$ . Таким образом, всегда будут выполняться условия:

$$C_{10} > C_{00}; \quad C_{01} > C_{11}. \quad (3.4)$$

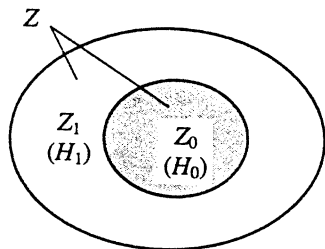


Рис. 3.1

Правило решения сводится к разбиению пространства  $Z$  наблюдения  $N$ -мерного вектора  $y$  на две части  $Z_0$  и  $Z_1$ , соответствующие принятию гипотез  $H_0$  и  $H_1$  (рис. 3.1).

Оптимизация правила решения сводится к такому выбору границ разбиения, определяющих решение, который минимизирует средний риск  $R$ .

Поскольку для области  $Z_0$  принимается гипотеза  $H_0$ , а для области  $Z_1$  — гипотеза  $H_1$ , то

$$P(H_0|H_0) = \int_{Z_0} p(y|u_0) dy, \quad (3.5)$$

$$P(H_1|H_0) = \int_{Z_1} p(y|u_0) dy, \quad (3.6)$$

$$P(H_1|H_1) = \int_{Z_1} p(y|u_1) dy, \quad (3.7)$$

$$P(H_0|H_1) = \int_{Z_0} p(y|u_1) dy, \quad (3.8)$$

где  $p(y|u_0)$  и  $p(y|u_1)$  — плотности условных вероятностей появления реализации  $y(t)$  при условии, что соответственно были переданы сигналы  $u_0$  и  $u_1$ . Интегралы от этих плотностей вероятности по областям  $Z_0$  и  $Z_1$  пространства реализаций  $y(t)$  дают вероятность того, что при этом реализация будет принадлежать данным областям и соответственно будут приняты гипотезы  $H_0$  и  $H_1$ .

Подставляя в (3.3) выражения (3.5)—(3.8), получаем

$$R = C_{00}P_0 \int_{Z_0} p(y|u_0) dy + C_{10}P_0 \int_{Z_1} p(y|u_0) dy + C_{11}P_1 \int_{Z_1} p(y|u_1) dy + C_{01}P_1 \int_{Z_0} p(y|u_1) dy. \quad (3.9)$$

Учитывая, что

$$\int_{Z_1} p(y|u_i) dy = 1 - \int_{Z_0} p(y|u_i) dy, \quad (3.10)$$

получаем

$$R = P_0C_{10} + P_1C_{11} + \int_{Z_0} [P_1(C_{01} - C_{11})p(y|u_1) - P_0(C_{10} - C_{00})p(y|u_0)] dy. \quad (3.11)$$

Первые два члена в (3.11) соответствуют фиксированным потерям, не зависящим от выбора границ области  $Z_0$ . Поэтому для минимизации  $R$  нужно выбрать область  $Z_0$  так, чтобы она включала все отрицательные (и только отрицательные) значения подынтегрального выражения в (3.11), т.е. из условия

$$P_1(C_{01} - C_{11})p(y|u_1) - P_0(C_{10} - C_{00})p(y|u_0) < 0. \quad (3.12)$$

Учитывая условие (3.4), неравенство (3.12) можно переписать в виде

$$\frac{p(y|u_1)}{p(y|u_0)} < \frac{P_0(C_{10} - C_{00})}{P_1(C_{01} - C_{11})}, \quad (3.13)$$

а правило решения в виде

$$\frac{p(y|u_1)}{p(y|u_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P_0(C_{10} - C_{00})}{P_1(C_{01} - C_{11})}. \quad (3.14)$$

Такая запись означает, что при знаке  $>$  принимается решение  $H_1$ , а при знаке  $<$  — решение  $H_0$ .

Левую часть неравенства (3.14) называют *отношением правдоподобия* и обозначают

$$\Lambda = \frac{p(y|u_1)}{p(y|u_0)}. \quad (3.15)$$

Смысл термина очевиден: более правдоподобна та из гипотез, которой соответствует большая плотность вероятности получения данной реализации  $y(t)$ . Каждой реализации случайной функции  $y(t)$  при любой размерности, используемой для задания дискретной выборки, соответствует

определенная величина отношения плотностей вероятностей в правой части (3.15). Поэтому отношение правдоподобия  $\Lambda(y)$  является одномерной случайной величиной.

Как будет видно из дальнейшего, отношение правдоподобия является фундаментальным соотношением, используемым не только в статистической теории обнаружения сигналов, но и в теории оценки параметров сигналов.

Правая часть (3.14) представляет фиксированную величину, зависящую только от априорного распределения вероятностей передачи сигналов и от выбора стоимостей верных и неверных решений. Она определяет порог решающей схемы

$$\eta = P_0 (C_{10} - C_{00}) / P_1 (C_{01} - C_{11}). \quad (3.16)$$

Таким образом, критерий Байеса сводится к критерию отношения правдоподобия

$$\Lambda(y) \underset{H_0}{\overset{H_1}{\geq}} \eta. \quad (3.17)$$

Важно отметить, что распределение априорных вероятностей и стоимостей влияет лишь на выбор численного значения порога, но не влияет на процедуру обработки данных, заключающуюся в вычислении  $\Lambda(y)$ . Поэтому уточнение стоимостей и априорных вероятностей, которые часто носят характер не очень строгих экспертных оценок на основе интуиции или предшествующего опыта, не влияет на решающую схему, а требует лишь перестройки уровня порога.

Так как обе части неравенства (3.17) положительны, то вместо отношения правдоподобия  $\Lambda(y)$  можно использовать в неравенстве его логарифм, являющийся монотонной функцией, и записать критерий отношения правдоподобия в форме

$$\ln \Lambda(y) \underset{H_0}{\overset{H_1}{\geq}} \ln \eta, \quad (3.18)$$

которая очень удобна, так как во многих случаях плотности вероятности  $p(y|u_i)$  являются показательными функциями.

Найдем выражение для  $\Lambda(y)$  в случае детерминированных сигналов  $u_i(t)$  и помехи  $n(t)$ , представляющей нормальный белый шум. Пусть спектр шума ограничен в пределах полосы  $\Delta f$ . В этом случае реализации шума  $n(t)$  в соответствии с теоремой Котельникова могут быть заданы дискретной выборкой  $(n_1, n_2, \dots, n_m)$  с интервалом между отсчетами  $\Delta t = 1/2\Delta f$ . Поскольку все эти отсчеты будут независимыми случайными величинами (см. § 2.8), многомерный нормальный закон распределения,

определяющий плотность вероятности  $w_m(n_1, n_2, \dots, n_m)$ , может быть представлен в виде (2.43) при  $\sigma_i = \sigma_m$ :

$$\begin{aligned} w_m(n_1, n_2, \dots, n_m) &= \frac{1}{(2\pi\sigma_m^2)^{m/2}} \prod_{k=1}^m \exp\left(-\frac{n_k^2}{2\sigma_m^2}\right) = \\ &= \frac{1}{(2\pi\sigma_m^2)^{m/2}} \exp\left(-\frac{1}{2\sigma_m^2} \sum_{k=1}^m n_k^2\right), \end{aligned} \quad (3.19)$$

где  $n_k$  — отсчеты белого шума  $n(t)$ , следующие через интервалы  $\Delta t = 1/2\Delta f$ ;  $m$  — общее число отсчетов на интервале  $T$  действия сигнала ( $m = T/\Delta t = 2T\Delta f$ ).

Плотность вероятности  $p(y|u_i)$  появления реализации входного сигнала  $y(t) = u_i(t) + n(t)$  соответствует плотности вероятности  $w_m[y(t) - u_i(t)]$  реализации шума  $n(t) = y(t) - u_i(t)$ . Дискретные отсчеты этой реализации шума  $n_k = y(t_k) - u_i(t_k) = \Delta y_{ik}$ , а соответствующая  $m$ -мерная плотность вероятности

$$\begin{aligned} p(y|u_i) &= w_m(\Delta y_{i1}, \Delta y_{i2}, \dots, \Delta y_{im}) = \\ &= \frac{1}{(2\pi\sigma_m^2)^{m/2}} \exp\left(-\frac{1}{2\sigma_m^2} \sum_{k=1}^m \Delta y_{ik}^2\right). \end{aligned} \quad (3.20)$$

Тогда

$$\begin{aligned} \Lambda(y) &= \frac{w_m(y - u_1)}{w_m(y - u_0)} = \exp\left\{-\frac{1}{2\sigma_m^2} \sum_{k=1}^m [\Delta y_{1k}^2 - \Delta y_{0k}^2]\right\} = \\ &= \exp\left(\frac{1}{N_0} \sum_{k=1}^m \{[y(t_k) - u_1(t_k)]^2 - [y(t_k) - u_0(t_k)]^2\} \Delta t\right). \end{aligned} \quad (3.21)$$

где  $N_0 = \sigma_m^2/\Delta f$  — спектральная плотность шума.

С учетом (3.21) критерий (3.18) для принятия решения о выборе гипотезы по известной реализации входного сигнала  $y(t)$  принимает вид

$$\begin{aligned} \ln \Lambda(y) &= -\frac{1}{N_0} \sum_{k=1}^m \{[y(t_k) - u_1(t_k)]^2 - \\ &- [y(t_k) - u_0(t_k)]^2\} \Delta t \underset{H_0}{\overset{H_1}{\geq}} \ln \eta. \end{aligned} \quad (3.22)$$

При достаточно большой ширине полосы  $\Delta f$  и соответственно малом  $\Delta t$  сумму можно заменить интегралом и записать критерий (3.22) в виде

$$-\frac{1}{N_0} \int_0^T \{ [y(t) - u_1(t)]^2 - [y(t) - u_0(t)]^2 \} dt \underset{H_0}{\geq} \underset{H_1}{\ln \eta}. \quad (3.23)$$

Сопоставляя левые части (3.23) и (1.20), легко видеть, что в случае обнаружения сигнала на фоне белого шума критерий Байеса сводится к сравнению расстояний в пространстве сигналов между принятой реализацией  $y(t)$  и передаваемыми сигналами  $u_1(t)$  и  $u_0(t)$ .

При использовании общего байесовского подхода к задаче обнаружения сигналов большую практическую трудность представляет выбор стоимостей  $C_{ij}$ , определяющих порог  $\eta$ . Поэтому при решении ряда практических задач используются критерии, представляющие модификации или частные формы критерия Байеса, позволяющие обойти эту трудность.

В последующих параграфах мы рассмотрим их применительно к конкретным задачам обнаружения сигналов.

### 3.3. Синхронный прием символов двоичного цифрового кода при действии белого шума. Критерий идеального наблюдателя

Пусть двоичный цифровой код используется для передачи значения некоторой случайной величины. Тогда ошибочные решения при приеме символа  $k$ -го разряда кода могут заключаться либо в приеме 1, когда был передан 0, либо в приеме 0, когда была передана 1. Абсолютное значение вызываемой этим ошибки в передаваемой величине в обоих случаях равно  $2^{k-1}$ , различие состоит лишь в знаке этой ошибки. Поэтому стоимость потерь при приеме символа  $k$ -го разряда в обоих случаях может быть принята одинаковой  $C_{10}^{(k)} = C_{01}^{(k)}$ . Правильному же приему символа можно приписать нулевые потери  $C_{11}^{(k)} = C_{00}^{(k)} = 0$ . Тогда минимизация среднего риска  $R$  сведется к минимизации среднего числа ошибок в приеме символов каждого разряда независимо от характера этих ошибок. Критерий минимума среднего числа ошибок был назван В. А. Котельниковым критерием "идеального наблюдателя" (здесь под "идеальным" понимается наблюдатель, беспристрастно фиксирующий любые ошибки). Иногда этот критерий называют еще критерием Котельникова.

Допустим теперь, что передаваемая цифровым кодом случайная величина с равной вероятностью может принимать любое значение от 0 до  $u_{\max}$ . В этом случае вероятности передачи в каждом из разрядов символов 0 и 1 одинаковы:  $P_0^{(k)} = P_1^{(k)} = 1/2$

При указанных условиях по формуле (3.16) получим  $\eta = 1$ , а  $\ln \eta = 0$  и критерий оптимального обнаружения символа любого разряда на фоне белого шума согласно (3.22) принимает вид:

$$\sum_{k=1}^m \{ [y(t_k) - u_0(t_k)]^2 - [y(t_k) - u_1(t_k)]^2 \} \Delta t \underset{H_0}{\geq} \underset{H_1}{0}. \quad (3.24)$$

Найдем теперь вероятности ошибочного приема (трансформации) символов двоичного кода  $P_{01}$  (прием 1 вместо 0) и  $P_{10}$  (прием 0 вместо 1). Рассмотрение будем вести для двух случаев передачи символов:

а) двоичный код с "пассивным" нулем: символ 1 передается сигналом  $u_1(t)$ , имеющим длительность  $T$  и энергию  $E$ , а символ 0 соответствует отсутствию этого сигнала на интервале длительностью  $T$  [ $u_0(t) = 0$ ];

б) двоичный код с "активным" нулем: символы 1 и 0 передаются различными сигналами  $u_1(t)$  и  $u_1(t)$ , имеющими одинаковые длительность  $T$  и энергию

$$\int_0^T u^2_1(t) dt = \int_0^T u^2_0(t) dt = E.$$

В первом случае критерий (3.24) принимает вид

$$2 \sum_{k=1}^m y(t_k) u_1(t_k) \Delta t - \sum_{k=1}^m u^2_1(t_k) \Delta t \underset{0}{\geq} 0$$

или

$$\sum_{k=1}^m y(t_k) u_1(t_k) \Delta t \underset{0}{\geq} \frac{1}{2} E, \quad (3.25)$$

где для обозначения гипотез вместо  $H_1$  и  $H_0$  используются принимаемые оценки 1 и 0.

Вероятность  $P_{01}$  трансформации нуля в единицу есть вероятность того, что случайная величина  $G_0$ , соответствующая левой части (3.25) при отсутствии сигнала  $u_1(t)$  в принятой реализации  $y(t)$ , т.е. при  $y(t_k) = n(t_k)$ , будет больше  $E/2$ .

Найдем закон распределения случайной величины

$$G_0 = \sum_{k=1}^m n(t_k) u_1(t_k) \Delta t. \quad (3.26)$$

Поскольку (3.26) представляет линейное преобразование белого шума, то  $G_0$  является центрированной случайной величиной с нормальным законом распределения

$$w_1(G_0) = \frac{1}{\sqrt{2\pi}\sigma_y} \exp\left(-\frac{G_0^2}{2\sigma_y^2}\right), \quad (3.27)$$

где  $\sigma_y^2$  — дисперсия случайной величины  $G_0$ ;

$$\sigma_y^2 = m_1(G_0^2) = m_1 \left\{ \left[ \sum_{k=1}^m n(t_k) u_1(t_k) \Delta t \right]^2 \right\} =$$

$$= \sum_{i=1}^m \sum_{k=1}^m m_1[n(t_i) n(t_k)] u_1(t_i) u_1(t_k) \Delta t^2 = \sum_{i=1}^m \sigma_m^2 u_1^2(t_i) \Delta t^2,$$

так как в силу некоррелированности случайных величин  $n(t_k)$ ,  $k=1, \dots, m$ , представляющих дискретную выборку реализации белого шума с интервалом между отсчетами  $\Delta t = 1/2\Delta f$  (см. § 2.8),

$$m_1[n(t_i) n(t_k)] = \begin{cases} \sigma_m^2 & \text{при } i = k, \\ 0 & \text{при } i \neq k. \end{cases}$$

Учитывая, что  $\sigma_m^2 \Delta t = \sigma_m^2 / 2\Delta f = N_0/2$ , получаем:

$$\sigma_y^2 = \frac{N_0}{2} \sum_{i=1}^m u_1^2(t_i) \Delta t = \frac{N_0 E}{2}.$$

Вероятность того, что случайная величина  $G_0$  с законом распределения (3.27) будет больше  $E/2$ , равна:

$$P_{01} = \int_{E/2}^{\infty} w_1(G_0) dG_0 = \frac{1}{\sqrt{2\pi}\sigma_y} \int_{E/2}^{\infty} \exp\left(-\frac{G_0^2}{2\sigma_y^2}\right) dG_0 =$$

$$= \frac{1}{\sqrt{2\pi}\sigma_y} \int_{-E/2}^{\infty} \exp\left(-\frac{G_0^2}{2\sigma_y^2}\right) dG_0 = F\left(-\frac{E}{2\sigma_y}\right) = F\left(-\sqrt{\frac{E}{2N_0}}\right), \quad (3.28)$$

где

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp\left(-\frac{t^2}{2}\right) dt$$

— интеграл вероятности.

Вероятность  $P_{10}$  трансформации единицы в нуль представляет вероятность того, что случайная величина  $G_1$  соответствующая левой части (3.25) при передаче сигнала  $u_1(t)$ , т.е. при  $y(t) = u_1(t) + n(t)$ , будет меньше  $E/2$ .

Случайная величина  $G_1$  может быть представлена в виде:

$$G_1 = \sum_{k=1}^m [u_1(t_k) + n(t_k)] u_1(t_k) \Delta t =$$

$$= \sum_{k=1}^m u_1^2(t_k) \Delta t + \sum_{k=1}^m n(t_k) u_1(t_k) \Delta t = E + G_0, \quad (3.29)$$

где  $G_0$  определяется формулой (3.26). Поэтому условие  $G_1 < E/2$  эквивалентно условию  $G_0 < -E/2$ , т.е.

$$P_{10} = \frac{1}{\sqrt{2\pi}\sigma_y} \int_{-\infty}^{-E/2} \exp\left(-\frac{G_0^2}{2\sigma_y^2}\right) dG_0 = F\left(-\sqrt{\frac{E}{2N_0}}\right) = P_{01}.$$

Таким образом, вероятности трансформации символов  $P_{10}$  и  $P_{01}$  одинаковы и определяются формулой (3.28). Равенство вероятностей  $P_{10}$  и  $P_{01}$  является прямым следствием критерия идеального наблюдателя, исходящего из равнозначности обеих ошибок.

При достаточно широкой полосе  $\Delta f$  (малом  $\Delta t$ ) сумму в левой части неравенства (3.25) можно заменить интегралом

$$\int_0^T y(t) u_1(t) dt \stackrel{1}{\geq} \frac{E}{2}, \quad (3.30)$$

по форме записи совпадающим с взаимной корреляционной функцией и имеющим название *корреляционного интеграла*.

Распознавание символов двоичного кода с пассивным нулем сводится, таким образом, к сравнению корреляционного интеграла с пороговым значением, равным половине энергии сигнала  $u_1(t)$ .

Тот же результат можно было бы получить несколько проще, используя фильтрующие свойства  $\delta$ -функции [1, 8], в предположении, что полоса пропускания приемника не ограничена и корреляционная функция шума есть  $\delta$ -функция. Однако мы сознательно отказались от этого допущения, а рассматривали физически реализуемый случай действия белого шума в конечной полосе  $\Delta f$ . Допущение о достаточно широкой полосе  $\Delta f$  делалось нами лишь для перехода от однократных сумм к интегралам, что не является принципиальным.

В случае передачи обоих символов сигналами  $u_1(t)$  и  $u_0(t)$  с одинаковой энергией  $E$  (двоичный код с "активным" нулем) критерий (3.24) при  $\eta = 1$  принимает вид:

$$\int_0^T y(t) [u_1(t) - u_0(t)] dt \stackrel{1}{\geq} 0 \quad (3.31)$$

или

$$\int_0^T y(t) u_1(t) dt \stackrel{1}{\geq} \int_0^T y(t) u_0(t) dt. \quad (3.32)$$

Таким образом, решение принимается в пользу того символа, для которого корреляционный интеграл принимает большее значение.

В этом случае  $P_{01}$  есть вероятность того, что при передаче сигнала  $u_0(t)$  левая часть (3.31) превысит нуль, а  $P_{10}$  — вероятность того, что при передаче сигнала  $u_1(t)$  левая часть (3.31) будет меньше нуля.

При передаче сигнала  $u_0(t)$  левая часть (3.31) будет представлять некоторую случайную величину  $G'_0$  равную:

$$G'_0 = \int_0^T [u_0(t) + n(t)] [u_1(t) - u_0(t)] dt = - \int_0^T u_0^2(t) dt + \int_0^T u_0(t) u_1(t) dt + \int_0^T n(t) [u_1(t) - u_0(t)] dt = -\lambda E + G', \quad (3.33)$$

где

$$\lambda = 1 - \frac{1}{E} \int_0^T u_1(t) u_0(t) dt, \quad (3.34)$$

$$G' = \int_0^T [u_1(t) - u_0(t)] n(t) dt. \quad (3.35)$$

Величину  $\lambda$  назовем *коэффициентом различимости символов*<sup>\*)</sup>. Если сигналы  $u_1(t)$  и  $u_0(t)$  тождественны (символы неразличимы), то  $\lambda = 0$ .

При передаче сигнала  $u_1(t)$  левая часть (3.31) будет представлять случайную величину  $G'_1$ , равную:

$$G'_1 = \int_0^T [u_1(t) + n(t)] [u_1(t) - u_0(t)] dt = \int_0^T u_1^2(t) dt - \int_0^T u_1(t) u_0(t) dt + \int_0^T n(t) [u_1(t) - u_0(t)] dt = \lambda E + G', \quad (3.36)$$

Из (3.33) и (3.36) следует, что условия  $G'_0 > 0$  и  $G'_1 < 0$  представляющие условия трансформации символов, эквивалентны условиям  $G' > \lambda E$  при трансформации 0 в 1 и  $G' < -\lambda E$  при трансформации 1 в 0.

Из сравнения (3.26) и (3.35) видим, что случайная величина  $G'$  получается из  $G_0$  заменой  $u_1(t)$  на  $[u_1(t) - u_0(t)]$ . Поэтому дисперсия нормальной случайной величины  $G'$  на основании (3.27) [при переходе от суммы к интегралу] равна

$$\begin{aligned} \sigma_y^2 &= \frac{N_0}{2} \int_0^T [u_1(t) - u_0(t)]^2 dt = \\ &= \frac{N_0}{2} \left[ \int_0^T u_1^2(t) dt - 2 \int_0^T u_1(t) u_0(t) dt + \int_0^T u_0^2(t) dt \right] = N_0 E \lambda. \end{aligned} \quad (3.37)$$

<sup>\*)</sup> Часто коэффициентом различимости называют величину

$$\frac{1}{E} \int_0^T u_1(t) u_0(t) dt,$$

однако предлагаемое обозначение представляется нам более удачным.

Вероятности трансформации символов по аналогии с (3.28) равны

$$P_{01} = P_{10} = F[-\lambda E / \sigma_y] = F(-\sqrt{\lambda E / N_0}). \quad (3.38)$$

Остановимся на геометрическом смысле коэффициента различимости сигналов. Воспользовавшись формулой (1.23), найдем расстояние между сигналами  $u_1(t)$  и  $u_0(t)$  при совмещении интервалов их существования  $T$ :

$$R_{u_1, u_0}^2 = \int_0^T [u_1(t) - u_0(t)]^2 dt = 2\lambda E. \quad (3.39)$$

Из (3.39) следует, что  $\lambda$  является мерой расстояния между сигналами. При  $\lambda = 0$  правая часть (3.39) обращается в нуль, что соответствует совпадению сигналов (сигналы неразличимы). Максимальное значение  $\lambda = 2$  получается при строго противофазных сигналах  $u_1(t)$  и  $u_0(t)$  и соответствует максимальному расстоянию между ними (максимальной различимости).

Из формулы (3.38) видно, что чем больше  $\lambda$  (лучше различимость сигналов), тем при том же отношении сигнала к шуму  $E/N_0$  меньше вероятность трансформации двоичных символов. При  $\lambda = 2$

$$P_{01} = P_{10} = F(-\sqrt{2E/N_0}), \quad (3.40)$$

а при  $\lambda = 0$

$$P_{01} = P_{10} = F(0) = 1/2,$$

что соответствует "гаданию вслепую".

Имеется простая связь между значением коэффициента различимости  $\lambda$  и выбором вида манипуляции несущей для передачи двоичных символов. При различении символов с использованием фазовой манипуляции несущей на  $\pi$ , когда сигналы  $u_1$  и  $u_0$  противофазны,  $\lambda = 2$ . Вероятность трансформации символов при этом минимальна и определяется формулой (3.38) при  $\lambda = 2$ . При частотной манипуляции, когда разность частот много больше величины  $1/T$  и сигналы  $u_1(t)$  и  $u_0(t)$  на интервале  $T$  можно считать ортогональными, коэффициент различимости  $\lambda = 1$ . При этом вероятность трансформации символов

$$P_{01} = P_{10} = F(-\sqrt{E/N_0}),$$

т.е. больше, чем при фазовой манипуляции, но меньше, чем для двоичного кода с пассивным нулем (3.28).



При амплитудной манипуляции  $0 \leq \lambda \leq 1$  и вероятность трансформации символов в общем случае еще больше. И только в частном случае импульсной амплитудной модуляции при условии, что импульсы, образующие сигналы  $u_1(t)$  и  $u_0(t)$ , не перекрываются, т.е. при ортогональности этих сигналов, амплитудная и частотная модуляции при равных энергиях сигналов эквивалентны.

Полученным результатам можно дать весьма наглядную геометрическую интерпретацию. Переходя к интегралу, левую часть (3.24) можно представить в виде

$$\int_0^T \{ [y(t) - u_0(t)]^2 - [y(t) - u_1(t)]^2 \} dt = R_{y,u_0}^2 - R_{y,u_1}^2,$$

где

$$R_{y,u_i}^2 = \int_0^T [y(t) - u_i(t)]^2 dt$$

в соответствии с (1.20) представляет квадрат расстояния между сигналами  $y(t)$  и  $u_i(t)$  в пространстве сигналов.

Таким образом, логарифм отношения правдоподобия при помехе в виде аддитивного белого шума пропорционален разности квадратов расстояний между входным сигналом  $y(t)$  и каждым из передаваемых сигналов  $u_1(t)$  и  $u_0(t)$ . Критерий же (3.24) означает, что выбирается тот из вариантов обнаруживаемых сигналов, которому соответствует меньшее расстояние до данной реализации сигнала  $y(t)$ . Этот критерий, исходящий из "равноправия" обнаруживаемых сигналов, совершенно естествен в условиях, когда оба варианта ошибочных решений одинаково опасны, а сами обнаруживаемые сигналы равновероятны.

Формула (3.38) для вероятностей трансформации символов двоичного цифрового кода с активным нулем с учетом (3.39) может быть записана в виде:

$$P_{10} = P_{01} = F\left(-\frac{0,5 R_{u_1,u_0}}{\sqrt{N_0/2}}\right). \quad (3.41)$$

Возвращаясь к формуле (3.28), определяющей вероятность трансформации символов двоичного цифрового кода с пассивным нулем, убеждаемся, что и она может быть представлена в виде (3.41). Действительно, для кода с пассивным нулем, когда  $u_0(t) \equiv 0$ , расстояние между сигналами  $R_{u_1,u_0} = \sqrt{E}$ , где  $E$  — энергия сигнала  $u_1(t)$  и, следовательно, аргумент функции  $F$  в (3.28) также равен

$$\frac{0,5 R_{u_1,u_0}}{\sqrt{N_0/2}}.$$

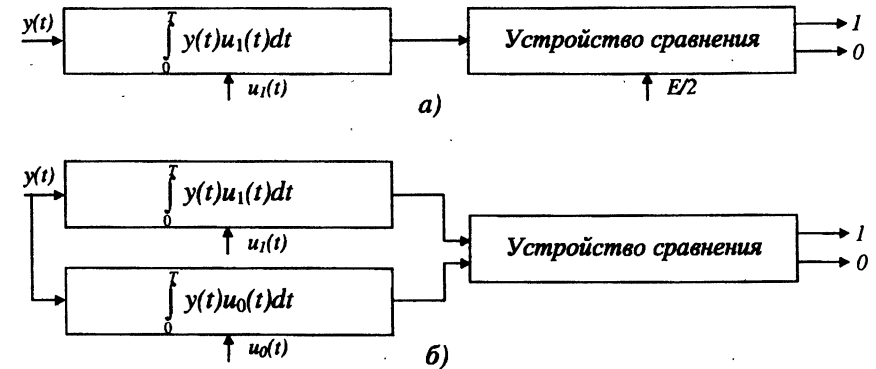


Рис. 3.2

Таким образом, при обнаружении бинарных равновероятных сигналов с использованием критерия идеального наблюдателя распознавание символов сводится к оценке близости отображающего их сигнала к принятой реализации  $y(t)$ . Мерой помехоустойчивости символов служит половина расстояния между обнаруживаемыми сигналами, что отражает тот факт, что трансформация символов наступает тогда, когда реализация  $y(t)$  из-за действия помехи отличается от переданного сигнала более чем на  $0,5 R_{u_1,u_0}$  и ближайшим к принятой реализации  $y(t)$  окажется не переданный, а другой сигнал.

Критерии принятия решений (3.30), (3.32) получены в предположении, что сигналы  $u_1(t)$  и  $u_0(t)$  совмещены в пространстве сигналов с полезной составляющей входного сигнала  $y(t)$  [это необходимо для выполнения условий  $y(t) - u(t) = n(t)$ , использовавшегося при выводе этих критериев]. В рассмотренных задачах приема символов двоичного кода выполнение этого условия обеспечивается синхронизацией, задающей начало приема очередного символа.

Решающая схема в данном случае может быть реализована с использованием коррелятора, вычисляющего корреляционные интегралы, входящие в критерии принятия решения (3.30), (3.32), и устройства сравнения (порогового устройства). При использовании двоичного кода с пассивным нулем значение корреляционного интеграла сравнивается с уровнем  $E/2$  (рис.3.2, а), а при использовании кода с активным нулем (рис.3.2, б) сравниваются значения корреляционных интегралов, входящих в (3.32).

### 3.4. Радиолокационное обнаружение.

#### Критерий Неймана—Пирсона

При радиолокационном обнаружении цели наряду с двумя благоприятными исходами, соответствующими правильно принятым решениям (цель обнаружена или установлен факт ее отсутствия), возможны два неблагоприятных исхода: принята гипотеза наличия цели, тогда как в действительности цели не было (ложная тревога) или принята гипотеза отсутствия цели, тогда как цель была (пропуск цели).

Выбор стоимостей потерь для этих двух ошибочных решений затруднен тем, что последствия ошибок в обоих случаях качественно разные: ложная тревога связана с лишними затратами физической и нервной энергии, а также материальных ресурсов, пропуск же цели — с большим экономическим ущербом. Поэтому определение порогового значения величины  $\eta$  по формуле (3.16) весьма затруднена. В качестве исходной величины для определения порогового значения  $\eta$  удобнее задаться допустимым значением вероятности ложной тревоги, что легче для экспертной оценки. В связи с этим в качестве критерия оптимизации режима радиолокационного обнаружения цели обычно используется критерий *Неймана—Пирсона*, согласно которому принимается гипотеза, соответствующая минимальной вероятности пропуска цели при заданном (экспертная оценка) допустимом значении вероятности ложной тревоги. Для того чтобы воспользоваться этим критерием, установим связь между пороговым значением отношения правдоподобия в формулах (3.17) или (3.18) и вероятностями ложной тревоги  $P_n$  и пропуска цели  $P_n$ .

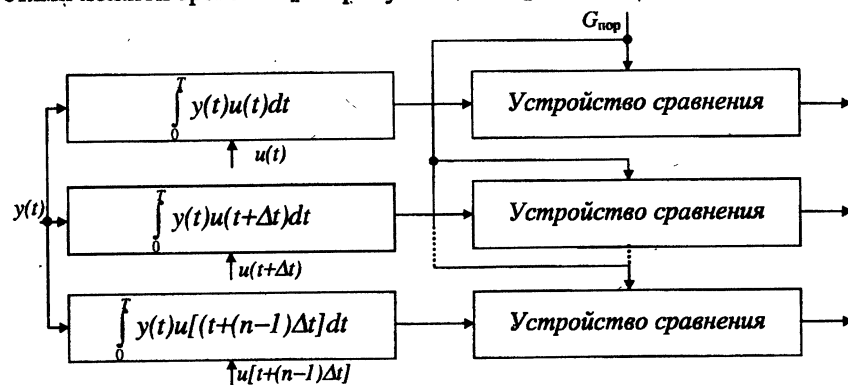


Рис. 3.3

Задача радиолокационного обнаружения сигнала известной формы сходна с уже рассмотренной задачей приема цифрового двоичного кода с

пассивным нулем: ложная тревога соответствует ошибочному приему символа 1, а пропуск цели — ошибочному приему символа 0. Однако задача радиолокационного обнаружения сигнала существенно отличается тем, что имеет место неопределенность времени задержки обнаруживаемого сигнала, тогда как при приеме символов двоичного кода предполагалось наличие синхронизации. Эту трудность можно преодолеть, если использовать многоканальный корреляционный приемник (рис. 3.3). В каналах этого приемника местные сигналы  $u(t)$  сдвинуты по времени на различные величины  $i\Delta t$ . Число каналов  $n$  выбирается так, чтобы величина  $n\Delta t$  перекрывала область неопределенности задержки обнаруживаемого сигнала. Тогда для одного из каналов всегда обеспечивается ошибка синхронизации местного и обнаруживаемого сигналов  $|\delta t| \leq \Delta t/2$ . Если число каналов выбрать так, чтобы для наиболее благоприятного канала погрешность синхронизации  $\delta t$  была достаточно малой и не оказывала существенного влияния на условия обнаружения, то для вычисления вероятностей  $P_n$  и  $P_n$  для этого канала можно использовать формулы для приема символов двоичного кода с пассивным нулем. Только в отличие от (3.30) пороговое значение корреляционного интеграла следует принимать равным не  $E/2$ , а некоторому уровню  $G_{\text{пор}}$  выбираемому исходя из допустимого значения вероятности ложной тревоги. Тогда, учитывая, что в соответствии с (3.26) и (3.29) случайные величины  $G_0$  и  $G_1$  отличаются лишь математическими ожиданиями, равными соответственно 0 и  $E$ , а также учитывая (3.27), для рассматриваемого канала получаем:

$$P_n = \int_{G_{\text{пор}}}^{\infty} w_1(G_0) dG_0 = \frac{1}{\sigma_y \sqrt{2\pi}} \int_{G_{\text{пор}}}^{\infty} \exp\left(-\frac{G_0^2}{2\sigma_y^2}\right) dG_0 = F\left(-\frac{G_{\text{пор}}}{\sqrt{N_0 E/2}}\right),$$

$$P_n = \int_{-\infty}^{G_{\text{пор}}} w_1(G_1) dG_1 = \frac{1}{\sqrt{2\pi}\sigma_y} \int_{-\infty}^{G_{\text{пор}}} \exp\left(-\frac{(G_1 - E)^2}{2\sigma_y^2}\right) dG_1 = F\left[\frac{(G_{\text{пор}} - E)}{\sqrt{N_0 E/2}}\right],$$

а критерий принятия решения принимает вид:

$$\int_0^T y(t) u(t) dt \underset{H_0}{\overset{H_1}{\geq}} G_{\text{пор}}.$$

Переходя к относительным величинам, перепишем полученные формулы в виде:

$$\frac{1}{E} \int_0^T y(t) u(t) dt \underset{H_0}{\overset{H_1}{\geq}} v_{\text{пор}}, \quad (3.42)$$

$$P_n = F(-v_{\text{пор}} \sqrt{2E/N_0}), \quad (3.43)$$

$$P_n = F[(v_{\text{пор}} - 1) \sqrt{2E/N_0}], \quad (3.44)$$

где  $v_{\text{пор}} = G_{\text{пор}}/E$ .

Произведем количественную оценку влияния погрешности синхронизации  $\delta\tau$  на характеристики обнаружения сигнала в предположении, что сохраняется критерий обнаружения (3.42), оптимальный при идеальной синхронизации. Очевидно, что погрешность синхронизации скажется лишь на случайной величине  $G_1(\delta\tau)$ , определяемой интегралом в левой части (3.42) в случае, когда сигнал  $u(t)$  передается. При этом входной сигнал

$$\eta(t) = u(t + \delta\tau) + n(t),$$

и с учетом стационарности шума  $n(t)$  можем записать

$$\begin{aligned} G_1(\delta\tau) &= \int_0^T [u(t + \delta\tau) + n(t + \delta\tau)] u(t) dt = \\ &= \int_0^T u(t + \delta\tau) u(t) dt + \int_0^T n(t) u(t) dt = E\rho(\delta\tau) + G_0, \end{aligned} \quad (3.45)$$

где

$$\rho(\delta\tau) = \frac{1}{E} \int_0^T u(t) u(t + \delta\tau) dt, \quad (3.46)$$

а

$$G_0 = \int_0^T n(t) u(t) dt$$

соответствует (3.26).

При отсутствии сигнала  $u(t)$  левая часть (3.42) представляет собой случайную величину  $G_0/E$ , не зависящую от  $\delta\tau$ .

Соответственно для вероятности ложной тревоги  $P_n$  и при наличии погрешности синхронизации остается справедливой формула (3.43). Влияние же погрешности синхронизации на вероятность пропуска цели  $P_n$ , как видно из сравнения (3.45) с (3.29), может быть учтено заменой в (3.44)  $E$  на величину  $E\rho(\delta\tau)$ :

$$P_n = F[(v_{\text{пор}} - 1) \sqrt{2E\rho(\delta\tau)/N_0}]. \quad (3.47)$$

Функцию  $\rho(\delta\tau)$ , совпадающую по своей форме представления с нормированной временной корреляционной функцией, назовем *нормированным корреляционным интегралом*. При  $\delta\tau=0$   $\rho(\delta\tau)=1$  и (3.47) переходит в (3.44). С увеличением  $\delta\tau$   $\rho(\delta\tau)$  уменьшается (рис. 3.4) и чувствительность корреляционного приемника к обнаруживаемому сигналу падает.

Для сигнала  $u(t)$  заданной формы всегда можно выбрать шаг  $\Delta t$  изменения задержки местного сигнала в многоканальном корреляционном приемнике, обеспечивающий соблюдение условия  $\rho(\delta\tau) \approx 1$  при  $\delta\tau \leq \Delta t/2$ , и для канала, удовлетворяющего этому условию, пользоваться формулой (3.44).

При анализе работы многоканального обнаружителя сигнала в целом следует учитывать, что решение об обнаружении цели принимается при принятии по критерию (3.42) гипотезы  $H_1$  хотя бы в одном канале. Поскольку ложная тревога имеет место в отсутствие сигнала, то в предположении

некоррелированного действия шума в отдельных каналах рассматриваемого обнаружителя суммарная вероятность ложной тревоги

$$P_{n\Sigma} = nP_n = nF(-v_{\text{пор}} \sqrt{2E/N_0}). \quad (3.48)$$

Вероятность пропуска цели для обнаружителя в целом, вообще говоря, меньше, чем для канала, обеспечивающего минимальную ошибку синхронизации, так как он в какой-то мере дублируется соседними каналами.

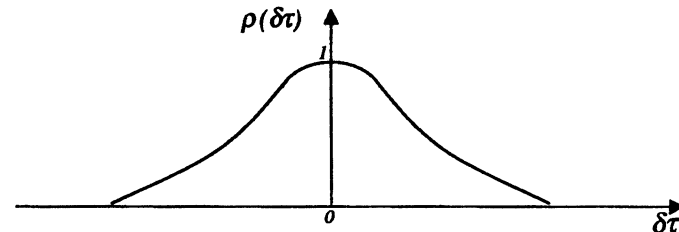


Рис. 3.4.

Однако при достаточно экономном выборе числа каналов, когда  $\Delta t$  соизмеримо с интервалом корреляции для принимаемого сигнала и величина  $\rho(i\Delta t)$  достаточно быстро падает с увеличением  $i$ , эффект дублирования невелик и в первом приближении можно принимать

$$P_{n\Sigma} \approx F(v_{\text{пор}} - 1) \sqrt{2E/N_0}. \quad (3.49)$$

Формула (3.48) позволяет определить  $v_{\text{пор}}$  для критерия обнаружения цели (3.42), по заданному допустимому значению  $P_{n\Sigma}$ , а (3.49) — соответствующую вероятность  $P_{n\Sigma}$ . Из этих формул видно, что с увеличением  $v_{\text{пор}}$  вероятность  $P_{n\Sigma}$  убывает, а вероятность  $P_n$  возрастает. Поэтому минимальное значение  $v_{\text{пор}}$  в критерии обнаружения цели (3.42), обеспечивающее требуемый уровень  $P_{n\Sigma}$  соответствует минимально возможной (при заданной вероятности  $P_{n\Sigma}$ ) вероятности пропуска цели, т.е. обеспечивает решение, соответствующее условиям критерия Неймана—Пирсона.

Из формул (3.38), (3.49) следует также, что вероятности  $P_n$  и  $P_n$  при оптимальном обнаружении зависят только от отношения энергии сигнала к спектральной плотности шума  $E/N_0$  и не зависят от формы сигнала.

Вместо многоканальной схемы, представленной на рис. 3.3, можно использовать последовательный поиск с перестройкой единственного канала и повторением посылок обнаруживаемого сигнала с энергией  $E$  через интервалы  $T$  (рис. 3.5). Схема рис. 3.5 может рассматриваться как

многоканальная с временным разделением каналов. Если можно пренебречь изменением задержки сигнала за время  $nT$ , то при одинаковых значениях  $T$  и  $E$  обе схемы эквивалентны и их характеристики определяются по тем же формулам. Упрощение аппаратуры при последовательном описке достигается ценой увеличения в  $n$  раз времени обнаружения.

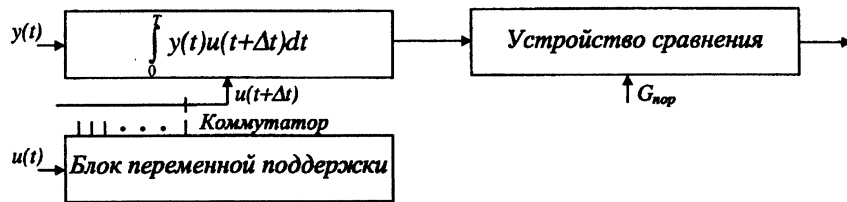


Рис. 3.5

Если обнаруживаемый сигнал, помимо неопределенности времени задержки, имеет достаточно большую область неопределенности доплеровского смещения частоты, то при его корреляционном приеме может потребоваться поиск в пространстве параметров сигнала как по временной задержки, так и по частоте (общее число каналов обнаружителя или время последовательного поиска соответственно возрастают).

Полученные результаты легко обобщаются на случай, когда обнаруживаемый сигнал имеет вид  $u(t; \alpha_1, \alpha_2, \dots, \alpha_k)$  и содержит случайные несущественные параметры  $\alpha_1, \alpha_2, \dots, \alpha_k$  не несущие информации, поиск по которым в пространстве сигнала не осуществляется. Многомерная функция распределения этих параметров  $w_k(\alpha_1, \alpha_2, \dots, \alpha_k)$  полагается известной. Поскольку принимаемое по выбранному критерию решение должно быть наилучшим в среднем, то нужно минимизировать усредненную по всем возможным значениям неизвестных параметров  $\alpha_1, \alpha_2, \dots, \alpha_k$  величину среднего риска  $R$ . В рассматриваемом случае обнаружения единственного сигнала  $u_1(t; \alpha_1, \alpha_2, \dots, \alpha_k)$  в формуле (3.11), определяющей величину  $R$  при бинарном обнаружении, следует положить  $u_0(t) \equiv 0$ . Осредняя в формуле (3.11) значение среднего риска  $R$  по всем возможным распределениям несущественных параметров обнаруживаемого сигнала, получаем

$$\bar{R} = P_0 C_{10} + P C_{11} + \int_{z_0} [P_1(C_{01} - C_{11}) \overline{p(y|u_1)} - P_0(C_{10} - C_{00}) p(y|u_0)] dy,$$

где

$$\overline{p(y|u_1)} = \int_0^T \dots \int_0^T p[y|u_1(t; \alpha_1, \dots, \alpha_k)] w_k(\alpha_1, \dots, \alpha_k) d\alpha_1 \dots d\alpha_k.$$

Таким образом все сводится к замене в (3.11) функции  $p(y|u_1)$  на  $\overline{p(y|u_1)}$  т.е. к замене отношения правдоподобия  $\Lambda(y)$  на величину

$$\bar{\Lambda}(y) = \frac{\overline{p(y|u_1)}}{p(y|u_0)} = \int_0^T \dots \int_0^T \Lambda(y; \alpha_1, \dots, \alpha_k) w_k(\alpha_1, \dots, \alpha_k) d\alpha_1 \dots d\alpha_k. \quad (3.50)$$

В качестве примера рассмотрим задачу обнаружения на фоне аддитивного белого шума узкополосного сигнала  $u(t, \varphi_0)$  со случайной начальной фазой. Представим сигнал  $u(t, \varphi_0)$  в виде колебания несущей частоты  $f_0$  с медленно меняющимися амплитудой и фазой и со случайной начальной фазой  $\varphi_0 = U(t) \cos[2\pi f_0 t + \varphi(t) + \varphi_0]$ . Отношение правдоподобия для фиксированного значения  $\varphi_0$  найдем, положив в (3.21)  $u_1 = u(t, \varphi_0)$  и  $u_0 = 0$ :

$$\begin{aligned} \Lambda(y) &= \exp \left( -\frac{1}{N_0} \int_0^T \{ [y(t) - U(t) \cos(2\pi f_0 t + \varphi(t) + \varphi_0)]^2 - y^2(t) \} dt \right) = \\ &= \exp \left( -\frac{E}{N_0} \right) \exp \left\{ \frac{2}{N_0} \int_0^T y(t) U(t) \cos[2\pi f_0 t + \varphi(t) + \varphi_0] dt \right\}, \end{aligned} \quad (3.51)$$

где

$$E = \int_0^T u^2(t, \varphi_0) dt$$

— энергия сигнала.

Осредненное значение отношения правдоподобия (3.51) в предположении равномерного распределения случайной фазы сигнала на интервале  $[0, 2\pi]$  равно

$$\bar{\Lambda}(y) = \exp \left( -\frac{E}{N_0} \right) \frac{1}{2\pi} \int_0^{2\pi} \exp \left\{ \frac{2}{N_0} \int_0^T y(t) U(t) \cos[2\pi f_0 t + \varphi(t) + \varphi_0] dt \right\} d\varphi_0. \quad (3.52)$$

преобразуем подынтегральную функцию в (3.52):

$$\begin{aligned} \exp \left\{ \frac{2}{N_0} \int_0^T y(t) U(t) \cos[2\pi f_0 t + \varphi(t) + \varphi_0] dt \right\} = \\ = \exp \left\{ \frac{2q_1}{N_0} \cos \varphi_0 - \frac{2q_2}{N_0} \sin \varphi_0 \right\} = \exp \left\{ \frac{2q}{N_0} \cos(\varphi_0 + \psi) \right\}, \end{aligned} \quad (3.53)$$

где

$$q_1 = \int_0^T y(t) U(t) \cos[2\pi f_0 t + \varphi(t)] dt = q \cos \psi; \quad (3.54)$$

$$q_2 = \int_0^T y(t) U(t) \sin[2\pi f_0 t + \varphi(t)] dt = q \sin \psi; \quad (3.55)$$

$$q = \sqrt{q_1^2 + q_2^2}; \quad \psi = \arctg(q_2/q_1). \quad (3.56)$$

Подставив (3.53) в (3.52), получим:

$$\bar{\Lambda}(y) = \exp \left( -\frac{E}{N_0} \right) \frac{1}{2\pi} \int_0^{2\pi} \exp \left[ \frac{2q}{N_0} \cos(\varphi_0 + \psi) \right] d\varphi_0 = \exp \left( -\frac{E}{N_0} \right) I_0 \left( \frac{2q}{N_0} \right), \quad (3.57)$$

где

$$I_0(x) = \frac{1}{2\pi} \int_0^{2\pi} \exp(x \cos \varphi) d\varphi = \frac{1}{2\pi} \int_0^{2\pi} \exp[-j(jx \cos \varphi)] d\varphi = J_0(jx) \quad (3.58)$$

— функция Бесселя  $J_0(jx)$  от мнимого аргумента  $jx$  [2, 34].

Критерий оптимального обнаружения (3.17) для отношения правдоподобия (3.57) принимает вид

$$I_0\left(\frac{2q}{N_0}\right) \underset{H_0}{\gtrsim} \eta e^{E/N_0}. \quad (3.59)$$

Решая неравенство (3.59) относительно величины  $q$  с учетом монотонно возрастающего характера функции  $I_0(2q/N_0)$  для положительных значений аргумента [34], получаем

$$q^2 \underset{H_0}{\gtrsim} \frac{CN_0}{2} = \eta_1 \quad (3.60)$$

где  $C$  — решение трансцендентного уравнения  $I_0(C) = \eta \exp(E/N_0)$ . Условие (3.60) можно переписать в виде

$$q^2 \underset{H_0}{\gtrsim} \eta_1^2.$$

т. е.

$$\left\{ \int_0^T y(t) U(t) \cos[\omega t + \varphi(t)] dt \right\}^2 + \left\{ \int_0^T y(t) U(t) \sin[\omega t + \varphi(t)] dt \right\}^2 \underset{H_0}{\gtrsim} \eta_1^2. \quad (3.61)$$

Интегралы в левой части (3.61) имеют вид корреляционных интегралов и отличаются только сдвигом на  $\pi/2$  начальной фазы высокочастотного заполнения используемых в них опорных сигналов.

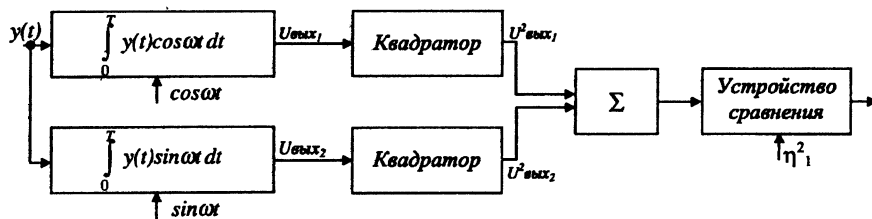


Рис. 3.6

В частном случае, когда обнаруживаемый сигнал представляет немодулированную посылку длительностью  $T$  (прямоугольный импульс) со случайной фазой высокочастотного сигнала

$$u(t, \varphi_0) = u_0 \cos(2\pi f_0 t + \varphi_0),$$

условие (3.61) принимает вид:

$$\left\{ \int_0^T [y(t) u_0(t) \cos 2\pi f_0 t] dt \right\}^2 + \left\{ \int_0^T [y(t) u_0(t) \sin 2\pi f_0 t] dt \right\}^2 \underset{H_0}{\gtrsim} \eta_1^2. \quad (3.62)$$

Интегралы в левой части (3.62) определяют так называемые квадратурные составляющие сигнала. Оптимальное обнаружение немодулированной посылки длительностью  $T$  со случайной фазой высокочастотного сигнала реализуется схемой квадратурного приема, изображенной на рис. 3.6.

### 3.5. Многоальтернативные задачи обнаружения

До сих пор мы рассматривали бинарные задачи обнаружения, соответствующие случаям, когда требуется принять одну из двух возможных гипотез о переданном сигнале (сигнал передавался или сигнал не передавался: передавался сигнал  $u_1$  или передавался сигнал  $u_0$ ). Рассмотрим теперь многоальтернативную задачу обнаружения<sup>\*)</sup>, когда может передаваться один из  $k$  заданных сигналов (в их числе может быть и нулевой сигнал), применительно к случаю, когда все ошибки обнаружения равноценны и требуется найти решение, дающее в среднем наименьшую вероятность ошибки (критерий идеального наблюдателя).

Перепишем применительно к этому случаю условие выбора решения (3.14) для бинарной задачи в виде:

$$\frac{P(y|u_1)}{P(y|u_0)} \underset{H_0}{\gtrsim} \frac{P_0}{P_1}$$

или

$$P_1 P(y|u_1) \underset{H_0}{\gtrsim} P_0 P(y|u_0).$$

Здесь  $P_0$  и  $P_1$  — априорные вероятности передачи сигналов  $u_0$  и  $u_1$  соответственно. Деля обе части этого неравенства на плотность вероятности  $P(y)$  данной реализации входного сигнала, в соответствии с (2.9) получаем условие выбора решения в виде:

$$P(H_1|y) \underset{H_0}{\gtrsim} P(H_0|y), \quad (3.63)$$

где  $P(H_i|y) = P_i p(y|u_i)/p(y)$  — апостериорная вероятность гипотезы  $H_i$  после получения данной реализации входного сигнала  $y(t)$ . Условие выбора (3.63) означает, таким образом, что выбирается гипотеза соответствующая максимальной апостериорной вероятности. Это условие выбора при использовании критерия идеального наблюдателя может быть распространено и на многоальтернативную задачу.

Таким образом, принятие решения по критерию идеального наблюдателя для многоальтернативной задачи обнаружения сводится к определению апостериорных вероятностей  $P(H_i|y)$  или пропорциональных им величин  $P_i p(y|u_i)$  для каждой из гипотез  $H_i$  и выбору той, для которой их значение максимально.

Рассмотрим конкретно случай приема сигналов с полностью известными параметрами на фоне белого шума. Для представления плотностей

<sup>\*)</sup> Часто задачу, когда нужно принять одну из нескольких возможных гипотез о переданном сигнале, называют задачей различения сигналов [2].

вероятности  $p(y|u_i)$  воспользуемся записью (3.20) плотности вероятности  $m$ -мерной выборки случайной функции  $y(t)$  с шириной спектра  $\Delta f$ :

$$p(y|u_i) = \frac{1}{(2\pi\sigma_{\text{ш}}^2)^{m/2}} \exp \left\{ -\frac{1}{2\sigma_{\text{ш}}^2} \sum_{k=1}^m [y(t_k) - u_i(t_k)]^2 \right\}.$$

Отбрасывая множитель  $1/(2\pi\sigma_{\text{ш}}^2)^{m/2}$ , не зависящий от выбора гипотезы  $H_i$  и не влияющий на определение максимальной из величин  $P_i p(y|u_i)$ , и заменяя сумму интегралом, как это делалось при выводе формулы (3.23), получаем

$$\max P_i p(y|u_i) \rightarrow \max P_i \exp \left\{ -\frac{1}{N_0} \int_0^T [y(t) - u_i(t)]^2 dt \right\},$$

и правило решения сводится к выбору гипотезы  $H_i$  для которой при данной реализации входного сигнала  $y(t)$  эта величина имеет максимальное значение. Заметим, что в показателе экспоненты стоит квадрат расстояния между сигналом  $u_i(t)$  и выходным сигналом, и второй множитель будет тем больше, чем меньше расстояние между этими сигналами.

### 3.6. Минимаксный критерий обнаружения сигналов

Рассмотрим случай, когда показатели потерь  $C_{ij}$  заданы, но неизвестны априорные вероятности обнаруживаемых сигналов  $P_1$  и  $P_0$  (рассмотрение будем вести применительно к бинарной задаче обнаружения).

Запишем критерий обнаружения Байеса (3.14) в виде

$$\Lambda(y) \underset{H_0}{\overset{H_1}{\geq}} \frac{(1 - P_1)(C_{10} - C_{00})}{P_1(C_{01} - C_{11})} = \eta(P_1). \quad (3.64)$$

Средний риск (3.11) можно записать в виде

$$\begin{aligned} R &= P_0 C_{10} + P_1 C_{11} + P_1(C_{01} - C_{11}) \int_{Z_0} p(y|u_1) dy - P_0(C_{10} - C_{00}) \times \\ &\times \left[ 1 - \int_{Z_0} p(y|u_0) dy \right] = P_1 C_{11} + (1 - P_1) C_{00} + P_1(C_{01} - C_{11}) P_{\text{н}} + \\ &+ (1 - P_1)(C_{10} - C_{00}) P_{\text{л}}, \end{aligned} \quad (3.65)$$

так как

$$\int_{Z_0} p(y|u_1) dy = P_{\text{н}}, \quad \int_{Z_0} p(y|u_0) dy = P_{\text{л}}.$$

Для каждого значения  $P_1$  можно рассчитать соответствующее значение порога  $\eta(P_1)$  в критерии (3.64) и определить значения  $P_{\text{н опт}}$  и  $P_{\text{л опт}}$  соответствующие условиям оптимального обнаружения по этому критерию. Подставляя  $P_{\text{н опт}}$  и  $P_{\text{л опт}}$  в (3.65), получаем функцию  $R_{\text{опт}}(P_1)$ , определяющую средний риск при оптимальном обнаружении в зависимости от априорных вероятностей обнаруживаемых бинарных сигналов (рис. 3.7).

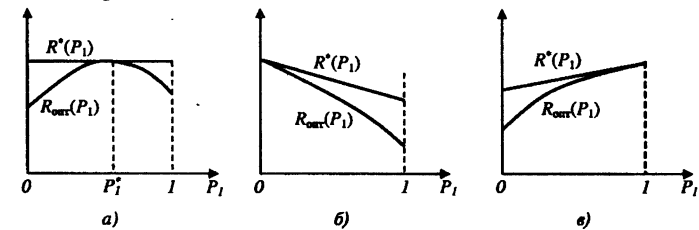


Рис. 3.7

Зафиксируем значение порога  $\eta^* = \eta(P_1^*)$  соответствующее априорной вероятности  $P_1^*$ , для которой  $R_{\text{опт}}(P_1^*)$  имеет максимальное значение. Если использовать это значение порога в критерии обнаружения (3.64), независимо от фактических значений априорной вероятности  $P_1$ , то вероятности  $P_{\text{н}}$  и  $P_{\text{л}}$ , зависящие только от выбора порога, будут постоянными для всех значений  $P_1$  и соответствующие значения среднего риска  $R^*(P_1)$  согласно (3.65) будут линейно зависеть от  $P_1$ . При  $P_1 = P_1^*$  когда порог  $\eta^*$  соответствует оптимальным условиям обнаружения,  $R^*(P_1^*) = R_{\text{опт}}(P_1^*)$ ; при всех других значениях  $P_1$  должно удовлетворяться условие  $R_{\text{опт}}(P_1) < R^*(P_1)$ , поскольку  $R_{\text{опт}}(P_1)$  — минимальный средний риск по смыслу критерия (3.64). Этим условиям удовлетворяет выпуклая кривая  $R_{\text{опт}}(P_1)$ , касающаяся прямой  $R^*(P_1)$  в точке максимума. Если максимум  $R_{\text{опт}}(P_1)$  лежит внутри интервала  $[0, 1]$  возможных значений  $P_1$ , то  $R^*(P_1)$  — горизонтальная прямая и средний риск при фиксированном пороге  $\eta^* = \eta(P_1^*)$  будет постоянным для любых значений априорной вероятности  $P_1$  появления сигнала  $u_1(t)$  (рис. 3.7, а). Если же экстремальное значение  $R_{\text{опт}}(P_1)$  достигается на краях интервала при  $P_1 = 0$  или  $P_1 = 1$  то  $R^*(P_1)$  может быть наклонной прямой с максимальным значением на соответствующем краю (рис. 3.7, б и в).

Таким образом, выбирая порог решающей схемы  $\eta^*$ , оптимальный для наиболее неблагоприятного значения  $P_1$  мы минимизируем максимальное значение риска на интервале возможных значений априорных вероятностей  $0 \leq P_1 \leq 1$ . Поэтому данный критерий принятия решения в условиях отсутствия данных об априорной вероятности  $P_1$  получил название *минимаксного*. При  $P_1 \neq P_1^*$  данный выбор фиксированного порога решающей схемы, естественно, не обеспечивает минимально возможного значения

среднего риска, однако при этом средний риск не будет превосходить минимально возможного его значения для наиболее неблагоприятного значения априорной вероятности  $P_1^*$ . В данном случае мы минимизируем не среднюю по всему диапазону возможных значений  $P_1$  величину  $R$ , а максимальную его величину на этом диапазоне. Используя минимаксный критерий, мы получаем гарантию, что при любом значении  $P_1$  не опустимся ниже уровня достоверности обнаружения, предельно достижимого для самых неблагоприятных сочетаний априорных вероятностей появления обнаруживаемых сигналов.

Найдем минимаксное решение задачи обнаружения сигнала  $u(t)$  на фоне аддитивного белого шума в предположении равноценности обеих видов ошибок (критерий идеального наблюдателя). Полагая  $C_{10} = C_{01} = C$ ,  $C_{11} = C_{00} = 0$ , получаем из (3.64) и (3.65)

$$\eta(P_1) = (1 - P_1)/P_1, \quad (3.66)$$

$$R(P_1) = C [P_1 P_n + (1 - P_1) P_n]. \quad (3.67)$$

Подставляя в (3.23) значения  $u_1(t) = u(t)$ ,  $u_0(t) = 0$  и  $\eta = (1 - P_1)/P_1$ , получаем:

$$\frac{1}{N_0} \left[ 2 \int_0^T y(t) u(t) dt - E \right] \gtrless_{H_0}^{\frac{H_1}{H_0}} \ln \frac{1 - P_1}{P_1}. \quad (3.68)$$

Приведем критерий обнаружения (3.68) к виду

$$\frac{1}{E} \int_0^T y(t) u(t) dt \gtrless_{H_0}^{\frac{H_1}{H_0}} \frac{N_0}{2E} \ln \frac{1 - P_1}{P_1} + \frac{1}{2}. \quad (3.69)$$

Вводя нормированное значение порога

$$v_{\text{пор}} = \frac{1}{2} + \frac{N_0}{2E} \ln \frac{1 - P_1}{P_1}$$

в (3.43) и (3.44), получаем:

$$P_n = F \left[ -\sqrt{\frac{E}{2N_0}} - \sqrt{\frac{N_0}{2E}} \ln \frac{1 - P_1}{P_1} \right], \quad (3.70)$$

$$P_n = F \left[ -\sqrt{\frac{E}{2N_0}} + \sqrt{\frac{N_0}{2E}} \ln \frac{1 - P_1}{P_1} \right]. \quad (3.71)$$

Используя (3.70) и (3.71), (3.67) можем записать в виде:

$$R(P_1) = C \left\{ P_1 F \left[ -\sqrt{\frac{E}{2N_0}} + \sqrt{\frac{N_0}{2E}} \ln \frac{1 - P_1}{P_1} \right] + (1 - P_1) F \left[ -\sqrt{\frac{E}{2N_0}} - \sqrt{\frac{N_0}{2E}} \ln \frac{1 - P_1}{P_1} \right] \right\}. \quad (3.72)$$

Можно показать, что в данном случае максимум  $R(P_1)$  соответствует условию равновероятных сигналов  $u_1(t) = u(t)$ , т.е. минимаксный критерий приводит к задаче, рассматривавшейся

в § 3.3. Вероятности ложной тревоги и пропуска сигналов для этих условий в соответствии с (3.28) равны:  $P_n = P_n = F(-\sqrt{E/2N_0})$ , а средний риск постоянен для всех значений  $P_1$ :

$$R^*(P_1) = C [P_1 F(-\sqrt{E/2N_0}) + (1 - P_1) F(-\sqrt{E/2N_0})] = CF(-\sqrt{E/2N_0}). \quad (3.73)$$

### 3.7. Особенности задачи селекции сигналов

До сих пор мы рассматривали обнаружение сигналов в условиях, когда они не налагаются друг на друга и их искажения обуславливаются только действием шума. Задача селекции сигналов возникает в случае, когда обнаруживаемые сигналы могут налагаться друг на друга. В этом случае входной сигнал  $\eta(t)$  может быть представлен в виде:

$$\eta(t) = A_1 u_1(t) + A_2 u_2(t) + n(t). \quad (3.74)$$

где коэффициенты  $A_1$  и  $A_2$  могут принимать дискретные случайные значения 1 (с вероятностями  $P_1$  и  $P_2$ ) или 0 (с вероятностями  $1 - P_1$  и  $1 - P_2$ ).

Особенность задачи селекции сигналов состоит в том, что помехой при обнаружении каждого из сигналов, помимо шума, может служить и налагающийся на него другой сигнал. Рассмотрим для определенности обнаружение сигнала  $u_1(t)$ . Решение о наличии сигнала  $u_1(t)$  ( $A_1 = 1$ ) или его отсутствии ( $A_1 = 0$ ) должно приниматься в результате сравнения с некоторым пороговым уровнем  $\eta$  отношения правдоподобия  $\Lambda(y)$ , которое в данном случае принимает вид:

$$\begin{aligned} \Lambda(y) &= \frac{w(y|A_1=1)}{w(y|A_2=0)} = \\ &= \frac{P_2 w(y|A_1=1, A_2=1) + (1 - P_2) w(y|A_1=1, A_2=0)}{P_2 w(y|A_1=0, A_2=1) + (1 - P_2) w(y|A_1=0, A_2=0)}, \end{aligned} \quad (3.75)$$

где  $w(y|A_1=1)$  — плотность вероятности получения реализации  $y(t)$  при наличии сигнала  $u_1(t)$  безотносительно к наличию или к отсутствию сигнала  $u_2(t)$ ;  $w(y|A_1=0)$  — плотность вероятности реализации  $y(t)$  при отсутствии сигнала  $u_1(t)$  безотносительно к наличию или к отсутствию сигнала  $u_2(t)$ ;  $w(y|A_1=1; A_2=1)$  — плотность вероятности реализации  $y(t)$  при наличии сигнала  $u_1$  и условии, что присутствует и сигнал  $u_2(t)$ ;  $w(y|A_1=1; A_2=0)$  — плотность вероятности реализации  $y(t)$  при наличии сигнала  $u_1(t)$  и условии, что сигнал  $u_2(t)$  отсутствует;  $w(y|A_1=0; A_2=1)$  — плотность вероятности реализации  $y(t)$  при отсутствии сигнала  $u_1(t)$  и наличии  $u_2(t)$ ;  $w(y|A_1=0; A_2=0)$  — плотность вероятности получения реализации  $y(t)$  при действии одного шума.

Выражая плотности вероятности, входящие в (3.75), через плотности вероятности соответствующих реализаций шума, получаем критерий для принятия решения о наличии или отсутствии сигнала  $u_1(t)$  в виде:

$$\frac{P_2 w_m(y - u_1 - u_2) + (1 - P_2) w_m(y - u_1)}{P_2 w_m(y - u_2) + (1 - P_2) w_m(y)} \gtrless_{A_1=0}^{A_1=1} \eta, \quad (3.76)$$

где  $\eta$  — пороговое значение отношения правдоподобия, определяемое формулой (3.16).

В случае белого шума левая часть (3.76), представляющая собой отношение правдоподобия  $\Lambda(y)$ , принимает вид

$$\begin{aligned} \Lambda(y) &= \frac{P_2 \exp \left\{ -\frac{1}{N_0} \int_0^T [y(t) - u_1(t) - u_2(t)]^2 dt \right\} +}{P_2 \exp \left\{ -\frac{1}{N_0} \int_0^T [y(t) - u_2(t)]^2 dt \right\} +} \rightarrow \\ &\quad + (1 - P_2) \exp \left\{ -\frac{1}{N_0} \int_0^T [y(t) - u_1(t)]^2 dt \right\} \\ &\rightarrow \frac{+ (1 - P_2) \exp \left\{ -\frac{1}{N_0} \int_0^T y^2(t) dt \right\}}{=} = \\ &= \frac{\exp \left\{ -\frac{1}{N_0} \int_0^T [y(t) - u_1(t)]^2 dt \right\}}{\exp \left\{ -\frac{1}{N_0} \int_0^T y^2(t) dt \right\}} \times \\ &\times \frac{P_2 \exp \left\{ -\frac{1}{N_0} \left[ -2 \int_0^T y(t) u_2(t) dt + 2E_{u_1, u_2} + E_2 \right] \right\} + (1 - P_2)}{P_2 \exp \left\{ -\frac{1}{N_0} \left[ -2 \int_0^T y(t) u_2(t) dt + E_2 \right] \right\} + (1 - P_2)}, \quad (3.77) \end{aligned}$$

где

$$E_2 = \int_0^T u_2^2(t) dt$$

— энергия сигнала  $u_2(t)$ ;

$$E_{u_1, u_2} = \int_0^T u_1(t) u_2(t) dt$$

— взаимная корреляция сигналов  $u_1(t)$  и  $u_2(t)$ .

Первая дробь в правой части (3.77) есть отношение правдоподобия для случая обнаружения сигнала  $u_1(t)$  в отсутствие других сигналов, а вторая характеризует влияние сигнала  $u_2(t)$  на условия обнаружения сигнала  $u_1(t)$ . Нетрудно видеть, что в случае  $E_{u_1, u_2} = 0$ , т.е. при ортогональности сигналов  $u_1(t)$  и  $u_2(t)$  на интервале  $[0, T]$ , вторая дробь в правой части (3.77) равна 1 и отношение правдоподобия  $\Lambda(y)$  отвечает условиям обнаружения сигнала  $u_1(t)$  на фоне белого шума. Полагая, что сигнал  $u_1(t)$  вне области  $[0, T]$  не существует, можем записать

$$E_{u_1, u_2} = \int_0^T u_1(t) u_2(t) dt = \int_{-\infty}^{\infty} u_1(t) u_2(t) dt.$$

Рассмотрим частный случай, когда  $u_1(t)$  и  $u_2(t)$  представляют результат отражения одного и того же радиолокационного сигнала от разных объектов и отличаются лишь временной задержкой  $\tau$  и доплеровским смещением частоты  $\mathcal{F}_d$ :

$$u_1(t) = u(t, 0, 0), \quad u_2(t) = u(t, \tau, \mathcal{F}_d).$$

В этом случае взаимное влияние сигналов на условия обнаружения каждого из них, т.е. условия разрешения радиолокационных объектов, определяются двумерной корреляционной функцией

$$B(\tau, F) = \int_{-\infty}^{\infty} u(t, 0, 0) u(t, \tau, \mathcal{F}_d) dt. \quad (3.78)$$

Чем острее пик корреляционной функции (3.78), тем при меньшем различии отраженных сигналов по временной задержке или доплеровскому смещению частоты будет ослабляться их взаимное влияние на условия обнаружения и соответственно будет обеспечиваться лучшее разрешение радиолокационных объектов.



## Глава 4. СТАТИСТИЧЕСКАЯ ТЕОРИЯ ОЦЕНКИ ПАРАМЕТРОВ СИГНАЛОВ

### 4.1. Байесовская оценка параметров сигнала. Функция риска

Пусть в полезном сигнале  $u(t, \alpha)$  нас интересует значение параметра  $\alpha$ , несущего информацию. Рассмотрим случай, когда на вход приемника поступает аддитивная смесь сигнала и шума.

$$y(t) = u(t, \alpha) + n(t) \quad (4.1)$$

Точное определение интересующего нас параметра  $\alpha$  в этих условиях невозможно. Задача заключается в определении так называемой оценки этого параметра  $\hat{\alpha} = \hat{\alpha}(y)$  по данной реализации входного сигнала  $u(t)$ , дающей наилучшее (в определенном смысле) приближение к действительному значению этого параметра.

Для определения критерия оптимальности оценки параметра  $\alpha$  так же, как и в задаче обнаружения, припишем каждому значению ошибки оценки  $\hat{\alpha} - \alpha$  определенные стоимости потерь (риска)  $R(\hat{\alpha} - \alpha)$ . Функцию  $R(\hat{\alpha} - \alpha)$  назовем *функцией риска*. Она является неотрицательной функцией ошибки оценки.

В качестве критерия оптимальности оценки естественно принять критерий минимума соответствующего ей среднего значения риска для всех возможных значений параметра  $\alpha$  и всех возможных реализаций смеси сигнала и шума  $y(t)$ . Этот критерий предложен Байесом и носит его имя. Его можно рассматривать как обобщение критерия Байеса, принятого при рассмотрении задачи обнаружения.

Полагая спектр реализаций  $y(t)$  ограниченным в пределах полосы  $\Delta f$ , можно задавать закон распределения этих реализаций соответствующим многомерным законам распределения дискретных выборок. Совместная плотность вероятности  $w_{\text{см}}(\alpha, y)$  принятой реализации  $y(t)$  и некоторого значения  $u(t, \alpha)$  параметра полезного сигнала  $u(t, \alpha)$  будет соответствовать плотности вероятности  $m$ -мерной выборки реализации шума  $n(t) = y(t) - u(t, \alpha)$ :

$$w_{\text{см}}(\alpha, y) = w_m[y(t_1) - u(t_1, \alpha), y(t_2) - u(t_2, \alpha), \dots, y(t_m) - u(t_m, \alpha)],$$

где  $t_1, t_2, \dots, t_m$  — дискретные моменты отсчета значений шума  $n(t)$ , следующие с интервалом  $\Delta t = 1/2\Delta f$ .

Величина среднего риска для выбранной функции риска  $R(\hat{\alpha} - \alpha)$  и известного закона распределения  $w_{\text{см}}(\alpha, y)$  равна

$$\bar{R} = \int_Z dy \int_{-\infty}^{\infty} R(\hat{\alpha} - \alpha) w_{\text{см}}(\alpha, y) d\alpha, \quad (4.2)$$

где  $Z$  — область  $m$ -мерного пространства, охватывающая дискретные выборки всех возможных реализаций  $y(t)$ . Поскольку внутренний интеграл в (4.2) положителен, то условие минимума величины  $\bar{R}$  можно заменить условием минимума внутреннего интеграла для каждой реализации  $y(t)$ , т.е. минимума условного риска, соответствующего частной реализации  $y(t)$ . Таким образом, требуется найти значение оценки параметра  $\hat{\alpha}$  удовлетворяющее условию:

$$\frac{d}{d\hat{\alpha}} \int_{-\infty}^{\infty} R(\hat{\alpha} - \alpha) w_{\text{см}}(\alpha, y) d\alpha = 0. \quad (4.3)$$

Очевидно, что получаемое по условию (4.3) значение оценки зависит от вида функции риска. Рассмотрим, во что обращается условие (4.3) для трех характерных форм функции риска:

1. Функция риска имеет вид параболы (рис. 4.1, а)

$$R(\hat{\alpha} - \alpha) = (\hat{\alpha} - \alpha)^2, \quad (4.4)$$

и критерий минимума среднего риска соответствует критерию минимума среднего квадрата ошибки оценки.

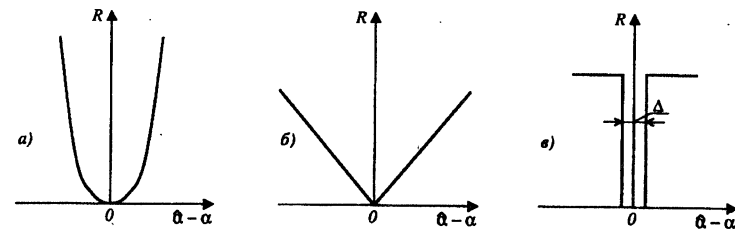


Рис. 4.1

2. Функция риска (рис. 4.1, б) имеет вид:

$$R(\hat{\alpha} - \alpha) = |\hat{\alpha} - \alpha|, \quad (4.5)$$

и критерий минимума среднего риска отвечает критерию минимума среднего значения модуля ошибки оценки.

3. Функция риска (рис. 4.1, в) имеет вид:

$$R(\hat{\alpha} - \alpha) = \begin{cases} 0 & \text{при } |\hat{\alpha} - \alpha| \leq \Delta/2, \\ R = \text{const} & \text{при } |\hat{\alpha} - \alpha| > \Delta/2. \end{cases} \quad (4.6)$$

Функция риска (4.6) характерна для систем, которые при ошибке оценки параметра, превосходящей величину  $\Delta/2$ , вообще не решают поставленную задачу, а при ошибке, не превышающей величину  $\Delta/2$ , обеспечивают полноценное решение поставленной задачи.

Для функции риска (4.4) условие (4.3) принимает вид

$$\begin{aligned} \frac{d}{d\hat{\alpha}} \int_{-\infty}^{\infty} w_{\text{см}}(\alpha, y) (\hat{\alpha} - \alpha)^2 d\alpha = \\ = 2\hat{\alpha} \int_{-\infty}^{\infty} w_{\text{см}}(\alpha, y) d\alpha - 2 \int_{-\infty}^{\infty} \alpha w_{\text{см}}(\alpha, y) d\alpha = 0. \end{aligned} \quad (4.7)$$

Используя (2.4), формулу (4.7) можно переписать в виде:

$$\hat{\alpha} w_m(y) \int_{-\infty}^{\infty} w_{\text{см}}(\alpha|y) d\alpha - w_m(y) \int_{-\infty}^{\infty} \alpha w_{\text{см}}(\alpha|y) d\alpha = 0,$$

где  $w_m(y)$  — плотность вероятности получения реализации  $y(t)$ , определяемой данной  $m$ -мерной выборкой ее отсчетов, при любых значениях параметра  $\alpha$ . Учитывая, что

$$\int_{-\infty}^{\infty} w_{\text{см}}(\alpha|y) d\alpha = 1,$$

получаем:

$$\hat{\alpha} = \int_{-\infty}^{\infty} \alpha w_{\text{см}}(\alpha|y) d\alpha. \quad (4.8)$$

Таким образом, оптимальная оценка параметра сигнала  $\hat{\alpha}$ , обеспечивающая минимум среднего квадрата отклонения оценки от истинного значения параметра, определяется как апостериорное математическое ожидание величины  $\alpha$ , соответствующее поступившей реализации  $y(t)$ , или что, то же самое, как абсцисса центра тяжести апостериорного закона распределения возможных значений параметра  $\alpha$ .

Для функции риска вида (4.5) условие (4.3) принимает вид:

$$\begin{aligned} \frac{d}{d\hat{\alpha}} \left[ \int_{-\infty}^{\infty} |\hat{\alpha} - \alpha| w_{\text{см}}(\alpha, y) d\alpha \right] = \frac{d}{d\hat{\alpha}} \left[ w_m(y) \int_{-\infty}^{\hat{\alpha}} (\hat{\alpha} - \alpha) \times \right. \\ \left. \times w_{\text{см}}(\alpha|y) d\alpha + w_m(y) \int_{\hat{\alpha}}^{\infty} (\alpha - \hat{\alpha}) w_{\text{см}}(\alpha|y) d\alpha \right] = 0 \end{aligned}$$

или

$$\int_{-\infty}^{\hat{\alpha}} w_{\text{см}}(\alpha|y) d\alpha = \int_{\hat{\alpha}}^{\infty} w_{\text{см}}(\alpha|y) d\alpha. \quad (4.9)$$

Таким образом, байесовская оценка параметра по критерию минимума среднего значения модуля ошибки соответствует абсциссе медианы апостериорного закона распределения  $w_{\text{см}}(\alpha|y)$ , делящей пополам площадь фигуры, отвечающей этому закону.

Во многих практических случаях, в частности при действии аддитивной нормальной помехи, функция  $w_{\text{см}}(\alpha|y)$  симметрична относительно ее максимума (рис. 4.2). В этом случае абсциссы центра тяжести и медианы апостериорного закона распределения  $w_{\text{см}}(\alpha|y)$  совпадают с его максимумом и оценка параметра для обоих рассмотренных выше видов функции риска соответствует максимуму апостериорной плотности вероятности  $w_{\text{см}}(\alpha|y)$ .

В случае, когда функция риска имеет вид (4.6), условие (4.3) соответствует условию минимума величины

$$\begin{aligned} w_m(y) \int_{-\infty}^{\infty} R(\hat{\alpha} - \alpha) w_{\text{см}}(\alpha|y) d\alpha = \\ = w_m(y) R \left[ 1 - \int_{\hat{\alpha}-\Delta/2}^{\hat{\alpha}+\Delta/2} w_{\text{см}}(\alpha|y) d\alpha \right]. \end{aligned} \quad (4.10)$$

Если при этом допустимая ошибка  $\Delta/2$  мала по сравнению с областью возможных значений параметра, то минимум выражения (4.10), отвечающий максимуму вычитаемого интеграла, обеспечивается при выборе оценки, соответствующей максимуму апостериорной плотности вероятности  $w_{\text{см}}(\alpha|y)$ .

Таким образом при некоторых дополнительных условиях (относительная малость величины  $\Delta$  при использовании функции риска вида (4.6) и симметрия апостериорного закона распределения  $w_{\text{см}}(\alpha|y)$  относительно его максимума), которые достаточно хорошо соблюдаются для многих практических задач, при всех трех существенно различных формах функции риска критерий Байеса свелся к общему критерию максимума апостериорной вероятности. Это свойство сравнительно слабой зависимости байесовской оценки от конкретного выбора функции риска при решении многих практических задач особенно важно, если учесть трудности строго обоснования формы функции риска.

Рассмотрим случай высокоточных измерений, когда область неопределенности значений параметра  $\alpha$  при известной реализации  $y(t)$ , определяемая апостериорным законом распределения  $w_{\text{см}}(\alpha|y)$ , много уже области возможных значений параметра по априорной информации (до измерений), определяемой распределением  $w(\alpha)$ . В этом случае в узкой области апостериорного распределения  $w_{\text{см}}(\alpha|y)$  можно принять априорное распределение  $w(\alpha) = \text{const}$ . Воспользуемся формулой Байеса (2.9), трансформировав ее применительно к случаю распределения значений непрерывных величин:

$$w_{\text{см}}(\alpha|y) = \frac{w(\alpha) w_{\text{см}}(y|\alpha)}{\int_{-\infty}^{\infty} w(\alpha) w_{\text{см}}(y|\alpha) d\alpha}. \quad (4.11)$$

Из (4.11) следует, что при  $w(\alpha) \equiv \text{const}$  апостериорная плотность вероятности  $w_{\text{см}}(\alpha|y)$  пропорциональна значению  $w_{\text{см}}(y|\alpha)$ . Функция  $w_{\text{см}}(y|\alpha)$ , представляющая плотность вероятности получения данной реализации  $y(t)$  при фиксированном значении  $\alpha$ , характеризует правдоподобность той или иной гипотезы о значении параметра  $\alpha$  при известной реализации  $y(t)$  и называется *функцией правдоподобия*. Критерий максимума апостериорной плотности вероятности при высокоточных измерениях сводится, таким образом, к критерию максимума функции правдоподобия.

Критерий максимума функции правдоподобия можно использовать также в случае, когда отсутствует априорная информация о законе распределения параметра, ибо при этом наиболее естественно принять это распределение равномерным.

## 4.2. Оптимальная оценка параметров сигнала при действии нормального белого шума

Рассмотрим случай высокоточных измерений параметра  $\alpha$ , полагая все остальные параметры сигнала известными. Сигнал в этом случае может быть записан в виде  $u(t, \alpha)$ . При действии белого нормального аддитивного шума апостериорная плотность вероятности  $w_{\text{см}}(\alpha|y)$  может приниматься симметричной относительно ее максимума. При этом оптимальная оценка параметра, как было показано, может производиться по критерию максимума функции правдоподобия. Введем вместо функции правдоподобия  $w_{\text{см}}(y|\alpha)$  отношение правдоподобия

$$\Lambda(y|\alpha) = \frac{w_{\text{см}}(y|\alpha)}{w_{\text{ш}}(y)}, \quad (4.12)$$

где  $w_{\text{ш}}(y)$  — плотность вероятности получения реализации  $y(t)$  при действии одного только шума. По форме записи зависимость вида  $w_{\text{ш}}(y)$ , приводящую в соответствие каждой реализации функции  $y(t)$  определенное число, следовало бы обозначать термином "функционал". Однако мы будем пользоваться термином "плотность вероятности", имея в виду, что реализации  $y(t)$  задаются дискретными  $m$ -мерными выборками и  $w_{\text{ш}}(y)$  обозначают плотность вероятности соответствующих выборок.

Поскольку знаменатель в (4.12) не зависит от  $\alpha$ , то максимум  $\Lambda(y|\alpha)$  по параметру  $\alpha$  совпадает с максимумом  $w_{\text{см}}(y|\alpha)$  и критерий максимума функции правдоподобия можно заменить критерием максимума отношения правдоподобия (4.12). Перепишем отношение правдоподобия (4.12) в виде:

$$\Lambda(y|\alpha) = \frac{w_{\text{см}}[y(t) - u(t, \alpha)]}{w_{\text{ш}}[y(t)]}, \quad (4.13)$$

учитывая, что плотность вероятности  $w_{\text{см}}(y|\alpha)$  реализации  $y(t) = u(t, \alpha) + n(t)$  эквивалентна плотности вероятности реализации шума  $n(t) = y(t) - u(t, \alpha)$ .

Нетрудно убедиться, что отношение правдоподобия (4.13) имеет точно такой же вид, как отношение правдоподобия (3.21) в бинарной задаче обнаружения при  $u_1(t) = u(t, \alpha)$ ,  $u_0(t) \equiv 0$ . Поэтому с учетом (3.21) — (3.23) получим

$$\ln \Lambda(y|\alpha) = - \frac{1}{N_0} \int_{-\infty}^{\infty} \{ [y(t) - u(t, \alpha)]^2 - y^2(t) \} dt. \quad (4.14)$$

То, что пределы интегрирования взяты от  $-\infty$  до  $+\infty$ , не нарушает общности формулы (4.14), поскольку фактические пределы интегрирования определяются областью существования сигнала  $u(t, \alpha)$ .

Преобразуя выражение (4.14), получаем

$$\begin{aligned} \ln \Lambda(y|\alpha) &= - \frac{1}{N_0} \left[ \int_{-\infty}^{\infty} u^2(t, \alpha) dt - 2 \int_{-\infty}^{\infty} u(t) u(t, \alpha) dt \right] = \\ &= \frac{2B(\alpha) - E(\alpha)}{N_0}, \end{aligned} \quad (4.15)$$

где

$$B(\alpha) = \int_{-\infty}^{\infty} y(t) u(t, \alpha) dt$$

называется корреляционным интегралом;

$$E(\alpha) = \int_{-\infty}^{\infty} u^2(t, \alpha) dt$$

— энергия сигнала при данном значении параметра  $\alpha$ .

Поскольку максимум функции (4.13) совпадает с максимумом ее логарифма, то оптимальную оценку параметра  $\alpha$  определим из условия:

$$\frac{d}{d\alpha} \left[ \frac{2B(\alpha)}{N_0} - \frac{E(\alpha)}{N_0} \right] = 0. \quad (4.16)$$

В частном случае, когда энергия сигнала не зависит от значения измеряемого параметра, получение оптимальной оценки параметра сводится к определению его значения  $\hat{\alpha}$  обеспечивающего максимум корреляционного интеграла  $B(\alpha)$ . Функциональная схема корреляционного приемника, реализующего оптимальную оценку параметра  $\alpha$  сигнала, приведена на рис. 4.3. В корреляторе 1 определяются значения корреляционного интеграла

$$B(\alpha_i) = \int_{iT}^{(i+1)T} y(t) u_0(t, \alpha_i) dt$$

на последовательных интервалах длительностью  $T$ , кратной периоду повторения сигнала. При этом в каждом цикле вычисления величины  $B(\alpha_i)$  параметр  $\alpha_i$  в опорном сигнале  $u_0(t, \alpha_i)$  меняется с достаточно малым шагом. В корреляторе 2 формируется величина  $B(\alpha_{i-1})$ , соответствующая на один шаг меньшему смещению параметра  $\alpha$  в опорном сигнале  $B(t, \alpha_{i-1})$ , поступающем с блока смещения по  $\alpha$  одновременно с сигналом  $B(t, \alpha_i)$ . Устройство сравнения фиксирует переход через нуль разности очередного значения корреляционного интеграла  $B(\alpha_i)$  и предшествующего его значения  $B(\alpha_{i-1})$ , (переход через максимум  $B(\alpha_i)$ ) и выдает сигнал на считывание соответствующего значения параметра  $\alpha$  в качестве оптимальной оценки  $\hat{\alpha}$ .

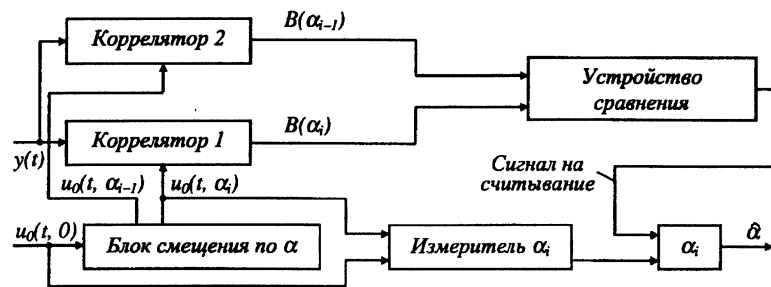


Рис. 4.3

Легко заметить большое сходство корреляционных приемников, работающих в режиме обнаружения сигнала (рис. 3.5) и в режиме оценки параметра (рис. 4.3). Отличие состоит в том, что в первом случае фиксируется факт превышения корреляционным интегралом некоторого порогового уровня при относительно грубой синхронизации местного и принимаемого сигнала, а во втором — значение параметра местного сигнала, при котором обеспечивается максимум корреляционного интеграла. Соответственно при поиске обнаруживаемого сигнала выбирается максимально допустимый шаг изменения параметров местного сигнала, не приводящий еще к критическому для достоверности обнаружения снижению максимального значения корреляционного интеграла (занижение этого шага приводило бы к неоправданному увеличению времени поиска). При оценке параметров шаг изменения параметров местного сигнала определяется требуемой точностью измерения. Отсюда следует, что режим обнаружения одновременно может служить и для грубой оценки параметров сигнала, позволяющей сузить область изменения параметров местного сигнала в режиме измерений.

Запишем корреляционный интеграл, получающийся при совместной обработке принятого сигнала  $y(t)$  и опорного сигнала  $u_0(t, \alpha)$  при некотором значении параметра  $\alpha$  в виде

$$B(\alpha) = \int_{-\infty}^{\infty} [u(t) + n(t)] u_0(t, \alpha) dt = B_0(\alpha) + B_n, \quad (4.17)$$

где

$$B_0(\alpha) = \int_{-\infty}^{\infty} u(t) u_0(t, \alpha) dt$$

— сигнальная составляющая корреляционного интеграла;

$$B_n = \int_{-\infty}^{\infty} n(t) u_0(t, \alpha) dt$$

— его шумовая составляющая.

Рассмотрим потенциальную точность измерения параметров сильного сигнала ( $E \gg N_0$ ) для случая, когда все параметры, кроме измеряемого, известны. Под потенциальной точностью измерения параметров полезного сигнала  $u(t)$  будем понимать тот предел точности, который при заданных форме и энергии сигнала и спектральной плотности шума не может быть улучшен при сколь угодно совершенной аппаратуре. Этот предел соответствует точности оптимальной оценки параметра сигнала при действии шума по критерию максимума отношения правдоподобия.

Найдем выражение для апостериорного закона распределения  $w_{\text{сш}}(\alpha|y)$  параметра  $\alpha$ . Как было показано, при высокоточных измерениях

он пропорционален функции правдоподобия  $w_{\text{см}}(y|\alpha)$ , которая по аналогии с (3.20) с учетом замены суммы интегралом, как в (3.21) — (3.23), с точностью до множителя, не зависящего от  $\alpha$ , определяется экспонентой

$$\exp \left\{ -\frac{1}{N_0} \int_{-\infty}^{\infty} [y(t) - u(t, \alpha)]^2 dt \right\}.$$

Поэтому, положив, что энергия сигнала

$$E = \int_{-\infty}^{\infty} u^2(t, \alpha) dt \equiv \int_{-\infty}^{\infty} y^2(t) dt$$

не зависит от значения параметра  $\alpha$ , получаем

$$w_{\text{см}}(\alpha|y) = k \exp \left\{ -\frac{1}{N_0} \left[ \int_{-\infty}^{\infty} y^2(t) dt - 2 \int_{-\infty}^{\infty} y(t) u(t, \alpha) dt + \int_{-\infty}^{\infty} u^2(t, \alpha) dt \right] \right\} = k \exp \left\{ \frac{1}{N_0} [2E - 2B(\alpha)] \right\} = k_1 \exp [2B(\alpha)/N_0]. \quad (4.18)$$

Разложим функцию  $B(\alpha)$  в окрестности оптимальной оценки  $\hat{\alpha}$ , соответствующей максимуму (вершине)  $B(\alpha)$ , в ряд Тэйлора, учитывая, что  $B'(\hat{\alpha}) = 0$ ,

$$B(\alpha) = B(\hat{\alpha}) + \frac{B''(\hat{\alpha})}{2} (\alpha - \hat{\alpha})^2 + \dots \quad (4.19)$$

Отбросив члены выше второго порядка малости, т.е. ограничившись аппроксимацией кривой  $B(\alpha)$  в окрестности точки  $\hat{\alpha}$  параболой, выражение (4.18) перепишем в виде:

$$w_{\text{см}}(\alpha|y) = k_1 \exp [2B(\hat{\alpha})/N_0] \exp \{ [B''(\hat{\alpha})/N_0] (\alpha - \hat{\alpha})^2 \} = c \exp \{ [B''(\hat{\alpha})/N_0] (\alpha - \hat{\alpha})^2 \}, \quad (4.20)$$

где  $c$  — некоторая константа.

Таким образом получаем, что апостериорный закон распределения  $w_{\text{см}}(\alpha|y)$  можно считать нормальным с дисперсией

$$\sigma_\alpha^2 = -N_0/2B''(\hat{\alpha}) \quad (4.21)$$

и средним значением  $\hat{\alpha}$ . [заметим, что правая часть (4.21) всегда положительна, так как по условию максимума  $B''(\hat{\alpha}) < 0$ .] Для сильного сигнала  $B_n \ll B_s(\alpha)$  и корреляционный интеграл  $B(\alpha)$  можно заменить его сигнальной составляющей  $B_s(\alpha)$ , приняв

$$B''(\hat{\alpha}) = B''_s(\alpha_0),$$

где  $\alpha_0$  — абсцисса точки максимума  $B_s(\alpha)$ .

Тогда для дисперсии оценки параметра получим следующее выражение:

$$\sigma_\alpha^2 = -\frac{1}{2B''_s(\alpha)/N_0} = \frac{1}{(2E/N_0) \rho''_s(\alpha)}, \quad (4.22)$$

где  $\rho_s(\alpha) = B_s(\alpha)/E$  — нормированная сигнальная составляющая корреляционного интеграла.

Закон распределения ошибки оценки соответственно имеет вид

$$w_{\text{см}}(\alpha|y) = \frac{1}{\sqrt{2\pi}\sigma_\alpha} e^{-(\alpha - \hat{\alpha})^2/2\sigma_\alpha^2}. \quad (4.23)$$

#### 4.3. Особенности оптимальной оценки параметров сигнала, зависящего от других случайных параметров, не подлежащих оценке

Рассмотрим теперь случай, когда сигнал имеет вид  $u(t, \alpha; \beta_1, \dots, \beta_k)$ , и зависит не только от оцениваемого параметра  $\alpha$ , но и от случайных параметров  $\beta_1, \dots, \beta_k$ , оценка которых не производится. Последние часто называют несущественными параметрами сигнала, а оцениваемый параметр  $\alpha$  — существенным.

В этом случае совместная плотность вероятности реализации входного сигнала  $y(t)$  и полезной его составляющей  $u(t, \alpha; \beta_1, \dots, \beta_k)$  зависит от сочетания всех параметров  $\alpha, \beta_1, \dots, \beta_k$ , определяющих полезный сигнал, и может быть записана в виде  $w_{\text{см}}(\alpha, \beta_1, \dots, \beta_k, y)$ .

Для получения функции среднего риска нужно провести осреднение условной функции риска по всем реализациям входного сигнала и всем возможным сочетаниям параметров  $\alpha, \beta_1, \dots, \beta_k$ , определяющих полезный сигнал:

$$\begin{aligned} \bar{R} &= \int_Z dy \int_{-\infty}^{\infty} R(\hat{\alpha} - \alpha) w_{\text{см}}(\alpha, \beta_1, \dots, \beta_k, y) d\alpha d\beta_1 \dots d\beta_k = \\ &= \int_Z dy \int_{-\infty}^{\infty} R(\hat{\alpha} - \alpha) \overline{w_{\text{см}}(\alpha, y)} d\alpha, \end{aligned} \quad (4.24)$$

где

$$\overline{w_{\text{см}}(\alpha, y)} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} w_{\text{см}}(\alpha, \beta_1, \dots, \beta_k, y) d\beta_1 \dots d\beta_k.$$

Таким образом, в данном случае формула для среднего риска отличается от (4.2) лишь тем, что вместо совместной плотности вероятности  $w_{\text{см}}(\alpha, y)$  используется осредненная по всем возможным значениям несущественных параметров совместная плотность вероятности  $\overline{w_{\text{см}}(\alpha, y)}$ . Поэтому и формулы для оценки параметра  $\alpha$  будут отличаться лишь тем,

что входящие в них вероятностные характеристики (апостериорная плотность вероятности, функция правдоподобия, отношение правдоподобия) предварительно осредняются по всем несущественным параметрам.

В качестве примера рассмотрим задачу оценки временной задержки  $\tau$  узкополосного сигнала

$$u(t, \tau, \varphi_0) = U(t - \tau) \cos [2\pi f_0(t - \tau) + \varphi(t - \tau) + \varphi_0], \quad (4.25)$$

где несущественным случайным параметром является начальная фаза  $\varphi_0$  высокочастотного заполнения сигнала.

Как и в предыдущем параграфе, будем рассматривать высокоточные измерения при действии помехи в виде аддитивного белого шума. В этом случае оптимальная оценка параметра  $\tau$  может производиться по критерию максимума отношения правдоподобия. Осредненное значение отношения правдоподобия для рассматриваемой задачи можно получить, воспользовавшись формулами (3.54)–(3.57), если заменить в них сигнал  $u(t, \varphi_0)$  на  $u(t, \tau, \varphi_0)$ . Тогда получим:

$$\Delta(y|\tau) = \exp(-E/N_0) I_0[2q(\tau)/N_0], \quad (4.26)$$

где  $E$  — энергия сигнала  $u(t, \tau, \varphi_0)$ , не зависящая от параметров  $\tau$  и  $\varphi_0$ ;

$$q(\tau) = \sqrt{q_1^2(\tau) + q_2^2(\tau)}, \quad (4.27)$$

$$q_1(\tau) = \int_{-\infty}^{\infty} y(t) U(t - \tau) \cos [2\pi f_0(t - \tau) + \varphi(t - \tau)] dt, \quad (4.28)$$

$$q_2(\tau) = \int_{-\infty}^{\infty} y(t) U(t - \tau) \sin [2\pi f_0(t - \tau) + \varphi(t - \tau)] dt. \quad (4.29)$$

Поскольку для положительных значений аргумента  $I_0[2q(\tau)/N_0]$  — монотонно возрастающая функция, максимум отношения правдоподобия (4.26) отвечает максимуму отношения  $2q(\tau)/N_0$ . При большом отношении сигнал-шум можно принять  $y(t) \equiv u(t, \tau, \varphi_0)$  и функцию  $q(\tau)$  заменить ее сигнальной составляющей

$$q_s(\tau) = \sqrt{q_{1s}^2(\tau) + q_{2s}^2(\tau)}, \quad (4.30)$$

где

$$\begin{aligned} q_{1s}(\tau) &= \int_{-\infty}^{\infty} U(t - \tau) \cos [2\pi f_0(t - \tau) + \varphi(t - \tau) + \varphi_0] U(t - \tau) \times \\ &\times \cos [2\pi f_0(t - \tau) + \varphi(t - \tau)] dt = \left\{ \int_{-\infty}^{\infty} U(t - \tau) \cos [2\pi f_0(t - \tau) + \right. \\ &\left. + \varphi(t - \tau)] U(t - \tau) \cos [2\pi f_0(t - \tau) + \varphi(t - \tau)] dt \right\} \cos \varphi_0 - \end{aligned}$$

$$\begin{aligned} &- \left\{ \int_{-\infty}^{\infty} U(t - \tau) \sin [2\pi f_0(t - \tau) + \varphi(t - \tau)] \times \right. \\ &\times U(t - \tau) \cos [2\pi f_0(t - \tau) + \varphi(t - \tau)] dt \left. \right\} \sin \varphi_0; \end{aligned} \quad (4.31)$$

$$\begin{aligned} q_{2s}(\tau) &= \left\{ \int_{-\infty}^{\infty} U(t - \tau) \cos [2\pi f_0(t - \tau) + \varphi(t - \tau)] \times \right. \\ &\times U(t - \tau) \sin [2\pi f_0(t - \tau) + \varphi(t - \tau)] dt \left. \right\} \cos \varphi_0 - \\ &- \left\{ \int_{-\infty}^{\infty} U(t - \tau) \sin [2\pi f_0(t - \tau) + \varphi(t - \tau)] \times \right. \\ &\times U(t - \tau) \sin [2\pi f_0(t - \tau) + \varphi(t - \tau)] dt \left. \right\} \sin \varphi_0; \end{aligned} \quad (4.32)$$

Обозначим первый интеграл в (4.31) через  $\mathcal{F}_{11}$  и второй интеграл в (4.32) — через  $\mathcal{F}_{22}$ . Эти интегралы представляют корреляционную функцию узкополосного сигнала (4.25) и в соответствии с (2.69) могут быть записаны в виде

$$\mathcal{F}_{11} = \mathcal{F}_{22} = B_{\text{инч}}(\tau - \tau') \cos 2\pi f_0(\tau - \tau'), \quad (4.33)$$

где  $B_{\text{инч}}(\tau - \tau')$  — корреляционная функция несущего информацию модулирующего сигнала для временного смещения  $\tau - \tau'$ .

Преобразуем два других интеграла, входящие в (4.31) и (4.32), представив временную задержку  $\tau$  в виде  $\tau = \tau'' - \tau'$ . Для узкополосного сигнала, фаза и амплитуда которого меняется медленно по сравнению с несущей частотой, можно принять:

$$U(t - \tau) \equiv U(t - \tau''); \quad \varphi(t - \tau) \equiv \varphi(t - \tau');$$

$$B_{\text{инч}}(\tau - \tau') \equiv B_{\text{инч}}(\tau'' - \tau').$$

Тогда второй интеграл в (4.31) можно привести к виду;

$$\mathcal{F}_{12} \equiv B_{\text{инч}}(\tau - \tau') \sin 2\pi f_0(\tau - \tau'), \quad (4.34)$$

а первый интеграл в (4.32)

$$\mathcal{F}_{21} \equiv -B_{\text{инч}}(\tau - \tau') \sin 2\pi f_0(\tau - \tau'). \quad (4.35)$$

Подставляя (4.33)–(4.35) вместо соответствующих интегралов в (4.31) и (4.32), а затем (4.31) и (4.32) — в (4.30), получим:

$$\begin{aligned} q_s(\tau) &= B_{\text{инч}}(\tau - \tau') \sqrt{(\cos 2\pi f_0(\tau - \tau') \cos \varphi_0 - \\ &\rightarrow - \sin 2\pi f_0(\tau - \tau') \sin \varphi_0)^2 + [\sin 2\pi f_0(\tau - \tau') \cos \varphi_0 + \\ &\rightarrow \end{aligned}$$

$$\rightarrow \overline{+ \cos 2\pi f_0(\tau - \tau') \sin \phi_0} = B_{\text{мн}}(\tau - \tau'). \quad (4.36)$$

Таким образом,  $q_s(t)$  представляет корреляционный интеграл для модулирующего сигнала. Поэтому оптимальная оценка величины запаздывания сигнала со случайной фазой  $\phi_0$ , отвечающая условию максимума отношения  $2q_s(\tau)/N_0 = 2B_{\text{мн}}(\tau - \tau')/N_0$  определяется из условия:

$$\frac{d}{d\tau} \left[ \frac{2B_{\text{мн}}(\tau - \tau')}{N_0} \right] = 0, \quad (4.37)$$

отличающееся от (4.16) при  $E(\alpha) = \text{const}$  только тем, что корреляционный интеграл определяется не для высокочастотного сигнала, а для несущего информацию модулирующего сигнала. Соответственно и в формуле (4.22) для определения дисперсии оценки в случае сигнала со случайной фазой вместо величины  $B''_s(t)$  должна использоваться вторая производная по параметру  $t$  от корреляционного интеграла модулирующего сигнала  $B'_{\text{мн}}(\tau - \tau')$ .

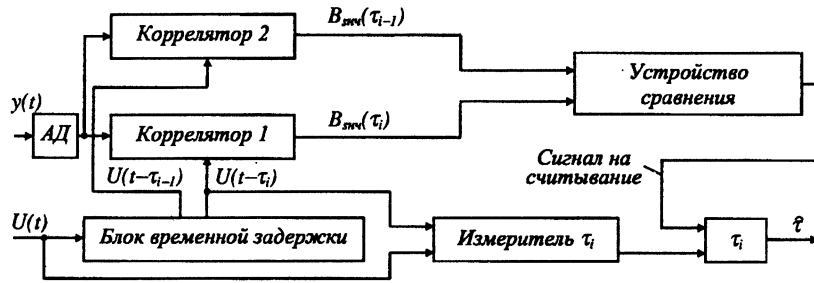


Рис. 4.4

Оптимальная схема корреляционного приемника в этом случае отличается от схемы, изображенной на рис. 4.3, лишь наличием детектора и использованием модулирующего сигнала в качестве опорного сигнала коррелятора. На рис. 4.4 эта схема приведена применительно к случаю амплитудно-модулированного сигнала. Введя обратную связь от устройства сравнения к блоку временной задержки, можно осуществить измерение задержки в режиме автоматического слежения за сигналом.

#### 4.4. Потенциальная точность измерения временной задержки и доплеровского сдвига частоты сигнала

При радиотехнических измерениях параметров движения объектов определение координат сводится к измерению временной задержки сигнала, а измерение составляющих скорости — к измерению доплеровского смещения частоты. Определим потенциальную точность измерения этих параметров.

Воспользуемся формулой (4.22) для определения потенциальной точности измерения задержки  $\tau$  сигнала  $u(t)$  произвольной формы с извест-

ными остальными параметрами. Сигнальная составляющая корреляционного интеграла  $B_s(t)$  в этом случае имеет вид:

$$B_s(\tau) = \int_{-\infty}^{\infty} u(t - \tau_0) u(t - \tau) dt \quad (4.38)$$

и соответственно

$$B''_s(\tau_0) = \int_{-\infty}^{\infty} u(t - \tau_0) u''(t - \tau_0) dt. \quad (4.39)$$

Интегрируя (4.39) по частям, получаем:

$$B''_s(\tau_0) = u(t - \tau_0) u'(t - \tau_0) \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} [u'(t - \tau_0)]^2 dt. \quad (4.40)$$

Первый член в правой части (4.40) равен нулю, поскольку реальные сигналы имеют конечную длительность и  $u(\infty) = u(-\infty) = 0$ . Поэтому из (4.40) получаем

$$\rho''_s(\tau_0) = \frac{B''_s(\tau_0)}{E} = - \frac{\int_{-\infty}^{\infty} [u'(t - \tau_0)]^2 dt}{\int_{-\infty}^{\infty} u^2(t - \tau_0) dt}. \quad (4.41)$$

На основании равенства Парсеваля можем записать

$$\int_{-\infty}^{\infty} u^2(t - \tau_0) dt = \int_{-\infty}^{\infty} |\dot{S}(f)|^2 df, \quad (4.42)$$

где  $\dot{S}(f)$  — комплексный спектр сигнала.

Модуль спектра производной сигнала  $u'(t - \tau_0)$  отличается от модуля спектра самого сигнала только множителем  $2\pi f$ , т.е. равен  $2\pi f |\dot{S}(f)|$ , и поэтому

$$\int_{-\infty}^{\infty} [u'(t - \tau_0)]^2 dt = 4\pi^2 \int_{-\infty}^{\infty} f^2 |\dot{S}(f)|^2 df. \quad (4.43)$$

На основании соотношений (4.42) и (4.43) формулу (4.41) можно записать в виде:

$$\rho''_s(\tau_0) = - \frac{4\pi^2 \int_{-\infty}^{\infty} f^2 |\dot{S}(f)|^2 df}{\int_{-\infty}^{\infty} |\dot{S}(f)|^2 df}. \quad (4.44)$$

Введем понятие эквивалентной ширины спектра сигнала  $\Delta f_s$ , определяемой соотношением

$$\Delta f^2 = \frac{4\pi^2 \int_{-\infty}^{\infty} f^2 |\dot{S}(f)|^2 df}{\int_{-\infty}^{\infty} |\dot{S}(f)|^2 df} \quad (4.45)$$

Тогда  $\rho''_s(\tau_0) = -\Delta f^2$ , и

$$\sigma^2_\tau = -\left[ \frac{2E}{N_0} \rho''(\tau_0) \right]^{-1} = \left[ \frac{2E}{N_0} \Delta f^2 \right]^{-1}. \quad (4.46)$$

Формула (4.46) имеет ясный физический смысл: чувствительность дальномерного канала к изменению задержки сигнала  $\tau$  при данном отношении  $E/N_0$  будет тем выше, чем более изрезанным будет "рельеф" сигнала  $u(t)$ , чем острее и выше будут отдельные пики этого рельефа. Вклад отдельных спектральных составляющих в рельеф сигнала зависит как от их частот, так и от амплитуды, что и учитывается в формуле для эквивалентной ширины спектра весовыми коэффициентами  $|\dot{S}(f)|^2$  при  $f^2$ . Значение  $\Delta f$  зависит от формы спектра сигнала, а при заданной его форме пропорционально ширине спектра. Например, для сигнала, имеющего в полосе  $\pm \Delta f/2$  равномерный сплошной спектр  $|\dot{S}(f)| = S_0$  получаем:

$$\Delta f^2 = \frac{4\pi^2 S_0^2 \int_{-\Delta f/2}^{\Delta f/2} f^2 df}{S_0^2 \int_{-\Delta f/2}^{\Delta f/2} df} = \frac{\pi^2 \Delta f^2}{3}.$$

Найдем потенциальную точность измерения доплеровского смещения  $\mathcal{F}_d$  частоты сигнала при известных остальных его параметрах. Сигнальную составляющую корреляционного интеграла  $B_s(\mathcal{F}_d)$  можно записать в виде

$$B_s(\mathcal{F}_d) = \int_{-\infty}^{\infty} u(t, \mathcal{F}_{d0}) u(t, \mathcal{F}_d) dt,$$

а ее вторую производную в виде

$$B''_s(\mathcal{F}_{d0}) = \int_{-\infty}^{\infty} u(t, \mathcal{F}_{d0}) u''_{\mathcal{F}_d}(t, \mathcal{F}_{d0}) dt. \quad (4.47)$$

Используем представление сигнала через его комплексный спектр  $\dot{S}(f)$ . Совместим начало отсчета времени с "центром масс" сигнала. Тогда смещение его средней частоты на величину  $\mathcal{F}_d$  отображается сдвигом по оси частот на ту же величину каждой из спектральных составляющих при неизменном значении их комплексного спектра. Поэтому можем записать

$$u(t, \mathcal{F}_d) = \int_{-\infty}^{\infty} \dot{S}(f) e^{j2\pi(f+\mathcal{F}_d)t} df, \quad (4.48)$$

где  $\dot{S}(f)$  — комплексный спектр сигнала  $u(t, 0)$ . После двукратного дифференцирования (4.48) по параметру  $\mathcal{F}_d$  получаем:

$$u''_{\mathcal{F}_d}(t, \mathcal{F}_d) = -4\pi^2 t^2 \int_{-\infty}^{\infty} \dot{S}(f) e^{j2\pi(f+\mathcal{F}_d)t} df. \quad (4.49)$$

На основании (4.47)–(4.49) имеем:

$$B''_s(\mathcal{F}_{d0}) = - \int_{-\infty}^{\infty} 4\pi^2 t^2 u(t, \mathcal{F}_{d0}) \left[ \int_{-\infty}^{\infty} \dot{S}(f) e^{j2\pi(f+\mathcal{F}_{d0})t} df \right] \times \\ \times dt = -4\pi^2 \int_{-\infty}^{\infty} t^2 u^2(t, \mathcal{F}_{d0}) dt,$$

откуда

$$\rho''_s(\mathcal{F}_{d0}) = - \frac{B''_s(\mathcal{F}_{d0})}{E} = - \frac{4\pi \int_{-\infty}^{\infty} t^2 u^2(t) dt}{\int_{-\infty}^{\infty} u^2(t) dt}.$$

Введем эквивалентную длительность сигнала  $T_s$ , определяемую соотношением:

$$T_s^2 = \frac{4\pi \int_{-\infty}^{\infty} t^2 u^2(t) dt}{\int_{-\infty}^{\infty} u^2(t) dt}, \quad (4.50)$$

где время  $t$  отсчитывается от середины ("центра масс") сигнала  $u(t)$ . Тогда получим

$$\sigma_{\mathcal{F}_d} = - \left[ \frac{2E}{N_0} \rho''_s(\mathcal{F}_{d0}) \right]^{-1} = \left[ \frac{2E}{N_0} T_s^2 \right]^{-1}. \quad (4.51)$$

Поясним физический смысл формул (4.50) и (4.51). Зависимость точности оценки частоты от длительности сигнала при фиксированной его форме очевидна: чем больше длительность сигнала, тем меньшие изменения частоты приводят к смещению фазы за время наблюдения, различимому на фоне шума. При определении эквивалентной длительности  $T_s$  (точнее, эквивалентного значения  $T_s^2$ ) для сигнала произвольной формы учитывается то обстоятельство, что вклад отдельных отсчетов сигнала в оценку частоты определяется удалением  $t$  момента отсчета от середины сигнала, характеризующим накапливающееся смещение фазы при данном смещении частоты. Поэтому величина  $T_s^2$  определяется по формуле (4.50) как средневзвешенное значение  $t^2$  с весовыми коэффициентами



$u^2(t)$ . В частности, для гармонической посылки длительностью  $\tau_c$  (прямоугольного радиоимпульса)  $T^2_c = \pi^2 \tau_c^2 / 3$ .

Из формул (4.46) и (4.51) следует, что при совмещении измерений координат и составляющих скорости с использованием простых сигналов, имеющих базу  $\Delta f T \equiv 1$ , мы сталкиваемся с противоречием: для повышения точности измерения координат следует стремиться к расширению спектра сигнала (что для простых сигналов означает уменьшение его длительности), а для повышения точности измерения составляющих скорости — к увеличению его длительности. Это противоречие является проявлением общего свойства сигналов, получившего название принципа неопределенности.

При использовании простых сигналов аналогичное противоречие возникает и при совмещении задач обнаружения и измерения координат. Дальность обнаружения, как было показано в гл. 3, зависит только от отношения энергии сигнала к спектральной плотности шума. При ограниченной пиковой мощности передатчика увеличение энергии сигнала требует увеличения его длительности, что приводит (для простого сигнала) к сужению во столько же раз его спектра и, в конечном счете, к уменьшению величины  $E \Delta f^2$ , а значит, и к ухудшению потенциальной точности измерения координат.

Оба эти противоречия разрешаются за счет применения сигналов сложной формы — так называемых широкополосных сигналов — с базой  $\Delta f T \gg 1$ .

#### 4.5. Совместная оценка фазы и частоты сигнала.

##### Теорема неопределенности

Важнейшими характеристиками измерительных каналов являются потенциальная точность и разрешающая способность. Рассмотрим их зависимость от характеристик сигнала, зафиксировав спектральную плотность шума.

Для потенциальной точности эта зависимость, как следует из § 4.3, достаточно полно характеризуется сигнальной составляющей  $B_s(\alpha)$  корреляционной функции сигнала по соответствующим его параметрам: по времени прихода — при измерении координат и по частоте — при измерении составляющих скорости. Чем больше вторая производная от корреляционной функции по измеряемому параметру  $B''_s(\alpha)$ , т.е. чем острее пик корреляционной функции, тем выше чувствительность системы к изменению параметра  $\alpha$ . Это хорошо видно при рассмотрении условий корреляционного приема, при котором обеспечивается оптимальная оценка параметров

сигналов известной формы: чем острее пик корреляционной функции, тем меньшие смещения опорного сигнала по измеряемому параметру ощутимы на фоне шума. При совместном измерении координат и составляющих скорости важна двумерная корреляционная функция  $B(\tau, \mathcal{F}_d)$ .

Разрешающая способность может рассматриваться либо в смысле обнаружения сигнала, либо в смысле измерения его параметров. Она определяется минимальным расстоянием между двумя сигналами в пространстве наблюдаемых параметров, при котором еще обеспечиваются установленные требования к характеристикам обнаружения или к точности измерения параметров каждого из сигналов. В § 3.6 при рассмотрении условий разрешения сигналов в смысле обнаружения было показано, что эти условия зависят от остроты пика двумерной корреляционной функции  $B(\tau, \mathcal{F}_d)$ , определяющей влияние соседнего сигнала на отношение правдоподобия для рассматриваемого сигнала. Поскольку отношение правдоподобия определяет и точность оценки параметров сигнала, то условия разрешения сигналов в смысле измерения будут также определяться остротой пика функции  $B(\tau, \mathcal{F}_d)$ .

Таким образом, и потенциальная точность, и разрешающая способность измерительных систем определяются одной и той же характеристикой сигнала — двумерной корреляционной функцией  $B(\tau, \mathcal{F}_d)$ . Эта общность требований к характеристикам сигнала легко объяснима близостью задач в обоих случаях. И определение точного положения сигнала в пространстве наблюдаемых параметров на фоне шума (потенциальная точность), и различение двух сигналов (разрешающая способность) сводятся к одному и тому же: способности различать близкие положения сигналов в пространстве наблюдаемых параметров. И хотя в первом случае речь идет о двух возможных положениях данного сигнала, а во втором — о двух сигналах, действующих одновременно, требования к форме сигнала остаются общими.

В общем случае при наличии амплитудной и угловой (фазовой или частотной) модуляции несущей сигнал можно записать в виде:

$$u(t) = U(t) \cos [2\pi f_0 t + \varphi(t) + \varphi_0] = \operatorname{Re} [\hat{U}(t) e^{j2\pi f_0 t}], \quad (4.52)$$

где  $\hat{U}(t) = U(t) e^{j\varphi(t)}$  — комплексная огибающая, определяющая закон амплитудной и угловой модуляции несущей;  $f_0$  — частота несущей.

Воспользовавшись комплексным представлением сигналов, запишем двумерную корреляционную функцию в следующем виде:

$$B(\tau, \mathcal{F}_d) = \int_{-\infty}^{\infty} \operatorname{Re} [\hat{U}(t) e^{j2\pi f_0 t}] \times$$

$$\times \operatorname{Re} [\dot{U}(t-\tau) e^{j2\pi(f_0-\mathcal{F}_d)(t-\tau)}] dt. \quad (4.53)$$

Преобразуем (4.53), воспользовавшись соотношением

$$\operatorname{Re} \dot{a} \operatorname{Re} \dot{b} = \operatorname{Re} [\dot{a}\dot{b}/2 + \dot{a}\dot{b}^*/2], \quad (4.54)$$

справедливость которого легко проверить (символом  $b^*$  обозначено комплексное число, сопряженное  $b$ ). Применяя к (4.53) преобразование вида (4.54) и учитывая, что действительная часть интеграла равна интегралу от действительной части подынтегральной функции, получаем:

$$B(\tau, \mathcal{F}_d) = \frac{1}{2} \operatorname{Re} \left[ e^{-j2\pi(f_0-\mathcal{F}_d)\tau} \int_{-\infty}^{\infty} \dot{U}(t) \dot{U}(t-\tau) \times \right. \\ \left. \times e^{j2\pi(f_0-\mathcal{F}_d)t} dt \right] + \frac{1}{2} \operatorname{Re} \left[ e^{j2\pi(f_0-\mathcal{F}_d)\tau} \int_{-\infty}^{\infty} \dot{U}(t) \dot{U}^*(t-\tau) e^{j2\pi\mathcal{F}_d t} dt \right]. \quad (4.55)$$

Обычно несущая частота  $f_0$  много больше ширины спектра сигнала. Поэтому, пренебрегая изменением комплексной огибающей  $\dot{U}(t)$  за период экспоненты  $e^{j2\pi(f_0-\mathcal{F}_d)t}$ , подынтегральную функцию в первом члене правой части (4.55) можно считать симметричной осциллирующей функцией и принять этот интеграл равным нулю. Тогда получим:

$$B(\tau, \mathcal{F}_d) = \operatorname{Re} \left[ \frac{1}{2} e^{j2\pi(f_0-\mathcal{F}_d)\tau} \int_{-\infty}^{\infty} \dot{U}(t) \dot{U}^*(t-\tau) \times \right. \\ \left. \times e^{j2\pi\mathcal{F}_d t} dt \right] = \operatorname{Re} [\dot{Z}(\tau, \mathcal{F}_d)].$$

Как и в случае однопараметрической корреляционной функции  $B(\tau)$  для узкополосного сигнала (см. § 2.9), функция  $\dot{Z}(\tau, \mathcal{F}_d)$  содержит два сомножителя: первый, не зависящий от формы сигнала, отображает изменение корреляционной функции с высокой частотой, и второй, определяемый формой сигнала, представляющий корреляционную функцию комплексной огибающей. Для характеристики разрешающей способности измерительного канала существен второй из этих множителей. Поэтому в качестве характеристики селективности сигнала можно использовать функцию

$$\psi(\tau, \mathcal{F}_d) = |\dot{Z}(\tau, \mathcal{F}_d)| = \\ = \frac{1}{2} \left| \int_{-\infty}^{\infty} \dot{U}(t) \dot{U}^*(t-\tau) e^{j2\pi\mathcal{F}_d t} dt \right|,$$

называемую *корреляционной функцией модуляции*. Нормируя ее, получаем:

$$\rho(\tau, \mathcal{F}_d) = \frac{\psi(\tau, \mathcal{F}_d)}{\psi(0, 0)} = \frac{\left| \int_{-\infty}^{\infty} \dot{U}(t) \dot{U}^*(t-\tau) e^{j2\pi\mathcal{F}_d t} dt \right|}{\left| \int_{-\infty}^{\infty} \dot{U}(t) \dot{U}^*(t) dt \right|}. \quad (4.56)$$

Однако для характеристики селективных свойств сигнала более удобно использовать функцию  $\rho^2(\tau, \mathcal{F}_d)$ . Функцию  $\rho^2(\tau, \mathcal{F}_d)$  называют *нормированной функцией неопределенности*. Поверхность, определяемая функцией  $\rho^2(\tau, \mathcal{F}_d)$ , получила название *поверхности неопределенности*, поскольку от ее формы зависит область, в пределах которой сигналы неразличимы при данном уровне шума. Объем, заключенный между этой поверхностью и координатной плоскостью  $\tau, \mathcal{F}_d$ , называется *телом неопределенности* (рис. 4.5, а).

Для упрощения графического изображения вместо тела неопределенности можно пользоваться диаграммами неопределенности, представляющими проекции на координатную плоскость  $\tau, \mathcal{F}_d$  горизонтальных сечений тела неопределенности на фиксированных уровнях (см. рис. 4.5, б).

Тело неопределенности обладает замечательным свойством инвариантности: *при любой форме сигнала полный объем тела неопределенности постоянен и равен единице*:

$$V_{\rho^2} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho^2(\tau, \mathcal{F}_d) d\tau d\mathcal{F}_d = 1. \quad (4.57)$$

Доказательство этого свойства, известного под названием *теоремы неопределенности*, мы приводить не будем (его можно найти, например, в [5]). Поскольку  $V_{\rho^2} = 1$  и  $\rho_{\max}(\tau, \mathcal{F}_d) = 1$ , то сужение пика тела неопределенности возможно только за счет соответствующего распределения остающегося объема на других участках области  $[\tau, \mathcal{F}_d]$  существования сигнала. Естественно, что любая деформация пика тела неопределенности отражает соответствующее изменение формы сигнала.

Оптимальный сигнал характеризуется:

1) максимальным сжатием основного пика тела неопределенности в направлении измеряемого параметра ( $\tau$  или  $\mathcal{F}_d$ ), что повышает потенциальную точность и селективность;

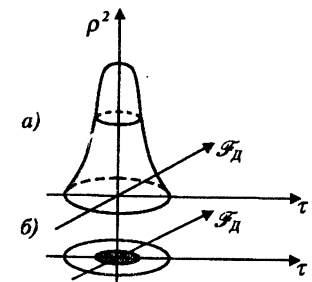


Рис. 4.5

2) минимальной высотой побочных пиков ("остатков") тела неопределенности в других частях области существования сигнала, что исключает неоднозначность измерений.

Из теоремы неопределенности следует, что принципиально возможны 3 пути сжатия пика тела неопределенности при соответствующем выборе формы сигнала.

1. Сжатие пика в направлении измеряемого параметра за счет расширения его в другом измерении. Этот путь эффективен только в случае, когда сигнал не используется для одновременного измерения координат и составляющих скорости. Примером может служить деформация пика тела неопределенности за счет изменения длительности импульсных посылок: при длительных посылках пик тела неопределенности сжат по координате  $\mathcal{F}_d$  и растянут по координате  $\tau$ , что благоприятно для измерения составляющих скорости; при очень коротких посылках пик, напротив, растянут по координате  $\mathcal{F}_d$ , но сжат по координате  $\tau$ , что благоприятно для измерения координат объекта, например дальности (на рис. 4.6 это проиллюстрировано соответствующими изменениями диаграмм неопределенности):

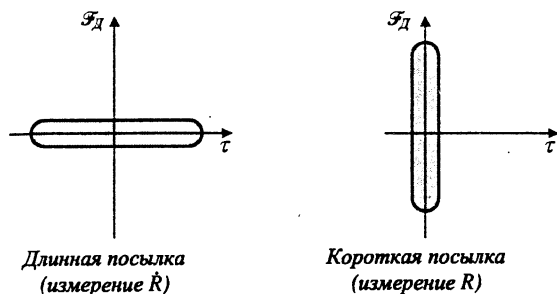


Рис. 4.6

2. Сжатие пиков за счет увеличения их числа в области существования сигнала (при периодическом сигнале), т.е. ценой ухудшения условий однозначности измерений. Примером может служить повышение точности измерения дальности фазовым методом с использованием гармонического модулирующего сигнала путем повышения частоты последнего. Повышение частоты модулирующего гармонического сигнала приводит к одновременному сжатию пиков тела неопределенности и уменьшению области однозначности измерения временной задержки, определяемой интервалом между пиками (рис. 4.7);

3. Сжатие пика за счет распределения остальной части тела неопределенности ("остатков") в виде тонкого по возможности равномерного слоя по всей остальной области  $[\tau, \mathcal{F}_d]$  существования сигнала (рис. 4.8).

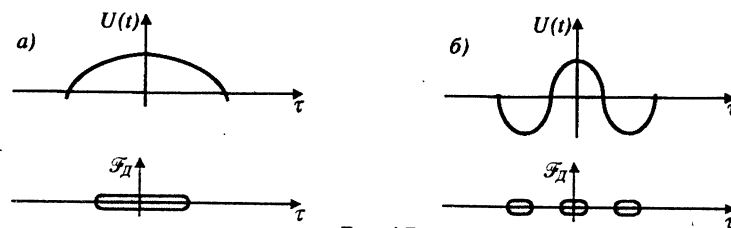


Рис. 4.7

Первый путь эффективен лишь для систем, предназначенных для измерения либо координат, либо составляющих скорости. Второй путь приемлем лишь в случае, когда область неопределенности измеряемого параметра по априорной информации достаточно мала. Третий является универсальным, поскольку обеспечивает возможность однозначного и точного измерения как координат, так и составляющих скорости. Поэтому для совмещенных систем, обеспечивающих одновременное измерение координат и составляющих скорости, оптимальным будет сигнал, обладающий телом неопределенности, изображенным на рис. 4.8. Такой сигнал должен прежде всего иметь возможно большую область существования  $[\tau, \mathcal{F}_d]$ , чтобы была возможность распределить основную часть объема тела неопределенности в виде достаточно тонкого слоя вне основного пика<sup>\*)</sup>. Поэтому оптимальные сигналы для совмещенных каналов измерения координат и составляющих скорости следует искать в классе так называемых широкополосных сигналов, для которых база сигнала  $\tau_c \mathcal{F}_c \gg 1$ , что обеспечивается при соответствующей форме сигнала (внутриимпульсной модуляции). К этому же выводу мы пришли и в предыдущем параграфе, отдельно рассматривая требования к форме сигнала для повышения потенциальной точности измерения координат и составляющих скорости.

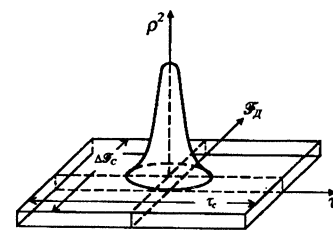


Рис. 4.8

#### 4.6. Широкополосные сигналы

Оптимальная обработка широкополосного сигнала должна сводиться к его сжатию по соответствующей координате в плоскости  $\tau, \mathcal{F}_d$ : по длительности сигнала — для измерения координат и по ширине спектра — для измерения скорости.

<sup>\*)</sup> Средний уровень "остатков" составит при этом примерно  $1/\tau_c \mathcal{F}_c$ .

Предельное сжатие широкополосного сигнала по длительности  $\tau_{\text{сж}}$ , определяющее потенциальную точность измерения координаты, зависит от ширины спектра исходного сигнала  $\mathcal{F}_c$ :

$$\tau_{\text{сж}} \equiv 1/\mathcal{F}_c = \tau_c/B, \quad (4.58)$$

где  $B = \tau_c \mathcal{F}_c$  — база сигнала. Очевидно, что при ширине спектра  $\mathcal{F}_c$  импульс короче  $\tau_{\text{сж}}$  не может быть сформирован.

Предельное сжатие сигнала по спектру  $\mathcal{F}_c$ , определяющее потенциальную точность измерения скорости, зависит от длительности исходного сигнала  $\tau_c$ :

$$\mathcal{F}_{\text{сж}} \equiv 1/\tau_c = \mathcal{F}_c/B. \quad (4.59)$$

При длительности исходного сигнала  $\tau_c$  более узкий спектр невозможен даже при полном снятии внутримпульсной модуляции.

Из формул (4.58) и (4.59) следует, что предельное сжатие широкополосного сигнала по любому измерению возможно в  $B$  раз.

В качестве примера широкополосных сигналов рассмотрим сигналы, манипулированные по фазе по закону бинарной псевдослучайной последовательности максимальной длительности, или сокращенно  $m$ -последовательностью. Эти сигналы называются фазоманипулированными псевдослучайными сигналами типа  $m$ -последовательности (сокращенно ФМПС).  $m$ -последовательность представляет собой набор  $N$  периодически повторяющихся символов  $d_i$ , каждый из которых может принимать одно из двух значений, например  $+1$  и  $-1$ . Между символами и законом фазовой манипуляции несущей устанавливается простая связь: индекс  $i$  определяет дискретные моменты времени  $t_i = i\tau_0$ , в которые может осуществляться манипуляция фазы на  $\pi$ ;  $d_i = +1$  отвечает соответственно фаз исходного и манипулированного сигналов на шаге, следующем за моментом  $t_i$ , а  $d_i = -1$  их противофазности на этом шаге. Таким образом, при каждой смене вида символов  $d_i$  фаза сигнала меняется на  $\pi$ .

Если сигнал несущей имеет вид  $u = u_0 \cos \omega t$ , то соответствующий ФМПС можно представить в виде:

$$u_{\text{пс}}(t) = u_0 d_i \cos \omega t = u_0 \cos [\omega t + (d_i - 1)\pi/2], \quad (4.60)$$

где  $i = [t/\tau_0]_{\text{цел}} -$  целая часть отношения  $t/\tau_0$ , задающая индекс символа  $d_i$ :

$$(d_i - 1) \frac{\pi}{2} = \begin{cases} 0 & \text{при } d_i = 1, \\ -\pi & \text{при } d_i = -1. \end{cases}$$

Для получения символов  $m$ -последовательности можно использовать закон формирования в виде произведения некоторого четного числа предшествующих символов [10]:

$$d_i = -d_{i-n} d_{i-l} \dots d_{i-k}, \quad (4.61)$$

где  $i, n, e, \dots, k$  — целые числа, причем  $i > n > e > \dots > l$ ; среди символов  $d_{i-1} \dots d_{i-n}$  хотя бы один равен  $+1$  (в противном случае все символы, формируемые по закону (4.61), равнялись бы  $-1$  и манипуляции фазы не было бы).

Нетрудно видеть, что предельная длительность периода повторения последовательности, формируемой по закону (4.61), равна  $N = 2^n - 1$  символов. Действительно, как следует из (4.61), в формировании  $i$ -го символа участвуют только символы из числа  $n$  ему предше-

ствующих. Поэтому как только повторятся исходные  $n$  символов, повторится и вся дальнейшая последовательность символов. Максимальная длительность периода последовательности определяется поэтому числом повторяющихся "слов" из  $n$  двоичных символов, начинающихся с каждого очередного символа этой последовательности, исключая слово, содержащее только символы  $-1$ . Максимальное число таких слов равно  $2^n - 1$ , что и определяет максимальную длительность периода последовательности. При соответствующем выборе числа символов и индексов  $n, e, \dots, k$  в законе (4.61) (см. § 9.3) оказывается возможным обеспечить формирование последовательности максимальной длительности ( $m$ -последовательности).

$m$ -последовательность обладает замечательным свойством: при посимвольном ее сравнении на интервале в один период с той же последовательностью, смещенной на  $l \neq pN$  элементов, число символов этих последовательностей, отличающихся по знаку, превышает число символов одинакового знака ровно на единицу [10]. Математически это свойство можно записать в виде:

$$\sum_{i=1}^N d_i d_{i-l} = -1. \quad (4.62)$$

При нечетном общем числе символов в периоде  $N = 2^n - 1$  лучшая их компенсация невозможна.

Найдем нормированную корреляционную функцию  $\rho(l\tau_0, \omega)$  ФМПС большой длительности ( $\tau_c \gg N\tau_0$ ) при отсутствии смещения по частоте. Будем считать, что шаг манипуляции фазы  $\tau_0$  кратен периоду несущей частоты и, следовательно, величина  $\omega(i-1)\tau_0$  кратна  $2\pi$ . Тогда воспользовавшись (4.56) и (4.60), для ФМПС бесконечной длительности получим:

$$\begin{aligned} \rho(l\tau_0, 0) &= \frac{\int_{-\infty}^{\infty} u(t) u(t - l\tau_0) dt}{\int_{-\infty}^{\infty} u^2(t) dt} = \frac{\int_0^{N\tau_0} u(t) u(t - l\tau_0) dt}{\int_0^{N\tau_0} u^2(t) dt} = \\ &= \frac{\sum_{i=1}^N u^2 d_i d_{i-l} \int_0^{N\tau_0} \sin^2 \omega t dt}{\sum_{i=1}^N u^2 d_i^2 \int_0^{N\tau_0} \sin^2 \omega t dt} = \frac{\sum_{i=1}^N d_i d_{i-l}}{N} = -\frac{1}{N}. \end{aligned} \quad (4.62)$$

Воспользовавшись периодичностью сигнала, мы заменили здесь отношение интегралов при бесконечных пределах интегрирования отношением интегралов, взятых за период  $m$ -последовательности, а последние заменили суммой интегралов за элементарные интервалы  $\tau_0$ . Заметим, что для сигналов ограниченной длительности это преобразование неправомерно, поскольку при этом область совместного существования сигналов  $u(t)$  и  $u(t - l\tau_0)$  не кратна периоду  $m$ -последовательности  $N\tau_0$ . Поэтому полученный результат справедлив лишь при  $\tau_c \gg N\tau_0$  (строго говоря, при  $\tau_c \rightarrow \infty$ ).

При  $l = pN$  сохраняется синфазность сигналов  $u(t)$  и  $u(t - l\tau_0)$  (в силу периодичности  $m$ -последовательности) и  $\rho(pN\tau_0, 0) = 1$ .

Нормированная корреляционная функция  $\rho(\tau, 0)$  для ФМПС большой длительности при отсутствии смещения по частоте имеет вид, приведенный на рис. 4.9. При корреляционном приеме такого сигнала происходит сжатие отрезков ФМПС длительностью  $N\tau_0$  в импульс длительностью  $\tau_0$ , т.е. в  $N$  раз. Нетрудно убедиться, что для ФМПС величина  $N$  представляет базу  $B$  сигнала длительностью в один период и полученное сжатие сигнала является, таким образом, предельным. Действительно, ширина спектра  $\mathcal{F}_0$  ФМПС определяется минимальным интервалом  $\tau_0$  между моментами манипуляции фазы и имеет порядок  $1/\tau_0$ . Поэтому при длительности сигнала  $\tau_c = N\tau_0$  его база имеет порядок  $B = \mathcal{F}_c \tau_c \equiv (1/\tau_0) N\tau_0 = N$ .

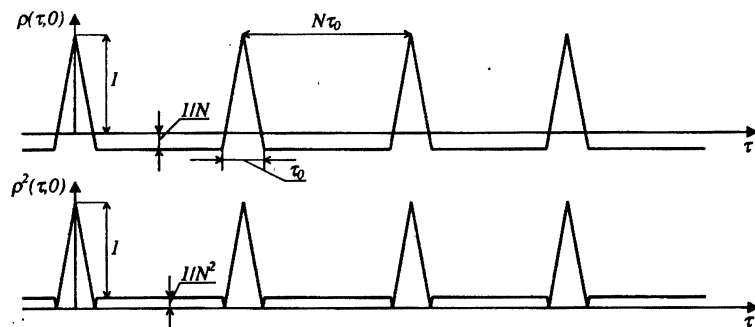


Рис. 4.9

Предельное сжатие сигнала по спектру до величины  $\mathcal{F}_{\text{сж}} \equiv 1/\tau_c$  может быть достигнуто снятием фазовой манипуляции. Для этого достаточно осуществить повторную манипуляцию фазы принятого сигнала точно по тому же закону  $m$ -последовательности, по которому осуществлялась манипуляция несущей при формировании ФМПС. Вид  $m$ -последовательности, используемой при формировании сигнала, предполагается известным на приемной стороне, а требуемая фаза ее определяется в результате измерения задержки сигнала  $\tau_0$  корреляционным методом. На рис. 4.9 приведен график сечения  $\mathcal{F} = 0$  нормированной корреляционной функции  $\rho(\tau, 0)$  и поверхности неопределенности  $\rho^2(\tau, 0)$  для ФМПС бесконечной длительности. Отношение уровня остатков к основным пикам тела неопределенности получается при этом равным  $1/N^2$ . При сечении для  $\mathcal{F} \neq 0$  уровень остатков будет несколько большим<sup>\*)</sup>. Тем не менее при достаточно большой базе уровень остатков ФМПС может быть обеспечен весьма малым.

ФМПС обладают рядом достоинств, обеспечивающих все более широкое их применение в современных радиотехнических системах:

- 1) близкая к оптимальной форма тела неопределенности, что делает их удобными в размещенных измерительных системах
- 2) относительная равномерность спектра в широкой полосе частот ("шумоподобность" сигнала), что делает его слабо различимым для средств радиоразводки на фоне шумов и обеспечивает высокую скрытность систем, работающих с ФМПС;
- 3) относительная простота генерации (см. § 9.3) приема и обработки сигналов;
- 4) возможность использования ФМПС, представляющих одну из разновидностей корректирующих кодов (см. § 9.3), одновременно и для измерений и для передачи информации.

<sup>\*)</sup> Средний уровень "остатков" во всей области существования сигнала близок к  $1/N$ .

## Глава 5

# ФИЛЬТРАЦИЯ СЛУЧАЙНЫХ СИГНАЛОВ

## 5.1. Общая характеристика задач фильтрации

Задачей фильтрации является получение из смеси полезного сигнала и шума оценки либо полезного сигнала в целом, либо меняющегося по случайному закону параметра этого сигнала, несущего полезную информацию. В первом случае критерием оптимальности фильтрации должно служить воспроизведение с минимальными искажениями формы (спектра) реализаций  $x_i(t)$  полезного случайного сигнала  $\xi(t)$ . Мерой качества фильтрации при этом может служить средний по множеству реализаций квадрат отклонения оценки сигнала  $\hat{\xi}(t)$ , получаемой на выходе фильтра, от истинной формы полезного сигнала  $\xi(t)$  (дисперсия оценки):

$$\sigma_{\xi}^2 = m_1\{[\xi(t) - \hat{\xi}(t)]^2\}, \quad (5.1)$$

а критерий оптимальности отвечает минимизации дисперсии оценки  $\sigma_{\xi}^2(t)$ .

В случае, когда задачей фильтрации является воспроизведение параметра сигнала, несущего информацию, критерии оптимальности фильтрации, естественно, будут другими, поскольку несоответствие формы входного и выходного сигналов подразумевается самой постановкой задачи. Во многих задачах выделения полезной информации таким критерием может служить получение максимального отношения сигнал/шум на выходе фильтра. Примером может служить случай передачи информации в цифровой форме, когда задача сводится к обеспечению наибольшей достоверности обнаружения сигналов, отображающих отдельные цифры. Как будет показано в § 5.4, этот же критерий обеспечивает оптимальность фильтрации такого параметра, как задержка сигнала.

Различают линейную и нелинейную фильтрацию. При линейной фильтрации сигналы претерпевают только линейные преобразования: усиление, суммирование, дифференцирование, интегрирование. При наличии обратных связей выходной сигнал также подвергается линейным преобразованиям. В общем виде выходной сигнал линейного фильтра  $y(t)$  определяется линейным дифференциальным уравнением вида:

$$a_n(t) \frac{d^n y}{dt^n} + a_{n-1}(t) \frac{d^{n-1} y}{dt^{n-1}} + \dots + a_1(t) \frac{dy}{dt} + a_0(t) y = f(t), \quad (5.2)$$

где  $f(t) = L[x(t)]$  — результат линейных преобразований входного сигнала  $x(t)$ , определяемых оператором  $L$ . Отсутствие в левой части уравнения (5.2) членов, отображающих интегральные линейные преобразования выходного сигнала, не отражается на его общности, поскольку дифференцированием обеих частей уравнения его всегда можно привести к форме (5.2). Переменные коэффициенты  $a_i(t)$  отвечают случаю линейной цепи с переменными параметрами. Линейная цепь с постоянными параметрами описывается линейным дифференциальным уравнением с постоянными коэффициентами

$$a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \dots + a_1 \frac{dy}{dt} + a_0 y = f(t). \quad (5.3)$$

При нелинейной фильтрации осуществляются нелинейные преобразования сигналов (перемножение, возведение в степень и др.). Выходной сигнал нелинейного фильтра в общем случае определяется нелинейным дифференциальным уравнением вида

$$a_n(y) \frac{d^n y}{dt^n} + a_{n-1}(y) \frac{d^{n-1} y}{dt^{n-1}} + \dots + a_1(y) \frac{dy}{dt} + a_0(y) y = f(t). \quad (5.4)$$

Основными свойствами линейных цепей, вытекающими из линейности описывающих их дифференциальных уравнений, являются линейная связь между изменениями входного сигнала  $x(t)$  и выходного сигнала  $y(t)$  и справедливость принципа суперпозиции, который можно сформулировать следующим образом: если входной сигнал представить в виде суммы элементарных сигналов

$$x(t) = \sum_{i=1}^m x_i(t),$$

то выходной сигнал будет представлять сумму соответствующих им элементарных выходных сигналов

$$y(t) = \sum_{i=1}^m y_i(t).$$

Эти два свойства, присущие только линейным цепям, резко упрощают как реализацию, так и математическое описание линейных фильтров, что и привело к выделению их в самостоятельный класс фильтров, получивших широкое применение. Естественно, ограничение используемых пре-

образований только линейными сказывается на возможностях фильтров: фильтр, оптимальный по заданному критерию в классе линейных фильтров, может существенно уступать по своим характеристикам оптимальному нелинейному фильтру. И только в частных случаях (см. § 5.3) поиск оптимальных фильтров без ограничений на вид преобразований приводит к линейным фильтрам. С другой стороны, имеются случаи, когда использование линейной фильтрации принципиально невозможно. Например, в случае, когда информационными параметрами являются фаза или частота сигнала, в силу нелинейной зависимости реализации сигнала от фильтруемого параметра может использоваться только нелинейная фильтрация. При этом оптимальными оказываются следующие фильтры (устройства фазовой или частотной автоподстройки частоты) [3].

Ограничимся рассмотрением лишь некоторых задач линейной фильтрации (читатели, интересующиеся задачами нелинейной фильтрации, могут обратиться к монографиям [1, 3, 4, 14, 38]).

## 5.2. Краткие сведения о линейных цепях и линейных преобразованиях сигналов

Принцип суперпозиции, справедливый для линейных цепей, позволяет определить их характеристики по отклику на некоторые сигналы стандартной формы (базисные функции), по которым может быть разложен входной сигнал. В качестве базисных функций при линейных преобразованиях сигналов обычно используются либо так называемые собственные функции, либо функции  $\delta(t)$  ( $\delta$ -импульсы).

Функция  $\theta(f, t)$  называется собственной, если она проходит через линейную цепь, не искажаясь по форме, а лишь меняя масштаб или претерпевая задержку по времени. Реакция цепи на собственную функцию  $\theta(f, t)$  определяется *собственным значением*  $H(f)$ , характеризующим масштаб ее линейного преобразования и задержку.

Для линейной цепи с постоянными параметрами собственными функциями являются экспоненциальные функции  $\{e^{j2\pi f t}\}$ , а собственными значениями — комплексные величины  $\dot{K}(f) = C(f) e^{j\varphi(f)}$ , определяющие отклик цепи на экспоненциальный комплексный сигнал  $e^{j2\pi f t}$ . Величина  $\dot{K}(f)$  называется *передаточной функцией* линейной цепи с постоянными коэффициентами, а величины  $C(f)$  и  $\varphi(f)$  — *амплитудно-частотной* и *фазочастотной* характеристиками цепи.

Если входной сигнал  $y(t)$  разложить в спектр по собственным функциям  $e^{j2\pi f t}$ , т.е. задать комплексным спектром  $\dot{S}(f)$ , то выходной сигнал  $y(t)$  будет равен сумме откликов линейной цепи на составляющие спектра

входного сигнала, т.е. будет определяться спектром  $\dot{K}(f) \dot{S}(f)$ . Таким образом, передаточная функция  $\dot{K}(f)$  полностью определяет линейное преобразование для случая цепи с постоянными параметрами.

Для линейных систем с переменными параметрами экспоненциальный комплексный сигнал уже не является собственной функцией. Отклик на него, представляющий сложное колебание, имеющее спектр частот, уже не может быть определен соответствующим значением передаточной функции. Поэтому применение методов гармонического спектрального анализа к таким системам оказывается неэффективным [14].

Помимо собственных функций, в качестве элементарных сигналов, используемых для характеристики (по отклику на них) линейных цепей, удобно применять простейшие временные сигналы —  $\delta$ -импульсы. Они

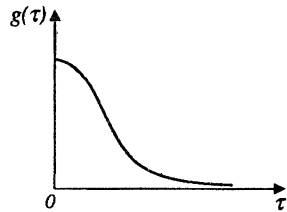


Рис. 5.1

удобны тем, что знание отклика на единственный сигнал  $\delta(t)$  полностью определяет линейную цепь, ибо спектральному разложению входного сигнала по  $\delta$ -функциям соответствует обычное временное представление  $x(t)$ .

Для цепей с переменными параметрами отклик на  $\delta$ -импульс зависит от конкретных значений момента

воздействия  $\delta$ -импульса на вход цепи и момента отсчета выходного сигнала. Функция  $g(t, t')$ , определяющая отклик в момент  $t$  на  $\delta$ -импульс, приложенный в момент  $t'$ , называется *импульсной переходной характеристикой* линейной цепи. Отклик на элементарный сигнал  $x(t) \delta(t')$ , представляющий спектральную составляющую разложения входного сигнала  $x(t)$  по  $\delta$ -функциям, соответственно равен  $x(t) g(t, t')$ . Используя принцип суперпозиции, получаем выходной сигнал как сумму откликов на воздействие элементарных входных сигналов, предшествовавших моменту  $t$ :

$$y(t) = \int_{-\infty}^t x(t') g(t, t') dt'. \quad (5.5)$$

Для линейных систем с постоянными параметрами отклик на  $\delta$ -импульс зависит только от интервала  $t - t' = \tau$  (рис. 5.1). В этом случае связь между входным и выходным сигналами принимает вид:

$$y(t) = \int_{-\infty}^t x(t') g(t - t') dt' = \int_0^{\infty} x(t - \tau) g(\tau) d\tau. \quad (5.6)$$

Найдем связь между импульсной переходной характеристикой  $g(\tau)$  и передаточной функцией  $\dot{K}(f)$  для линейной цепи с постоянными параметрами.

рами. Для этого, воспользовавшись формулой (5.6), найдем отклик линейной цепи с импульсной переходной характеристикой  $g(\tau)$  на входной экспоненциальный комплексный сигнал  $e^{j\omega t}$ :

$$\int_0^{\infty} e^{j2\pi f(t-\tau)} g(\tau) d\tau = e^{j2\pi f t} \int_0^{\infty} e^{-j2\pi f \tau} g(\tau) d\tau. \quad (5.7)$$

Сравнивая (5.7) со значением этого же отклика  $\dot{K}(f) e^{j2\pi f t}$ , определенным через передаточную функцию  $\dot{K}(f)$ , получаем:

$$\dot{K}(f) = \int_0^{\infty} e^{-j2\pi f \tau} g(\tau) d\tau = \int_{-\infty}^{\infty} e^{-j2\pi f \tau} g(\tau) d\tau, \quad (5.8)$$

так как для физически реализуемых цепей  $g(\tau) = 0$  при  $\tau < 0$  (в момент  $t$  не может существовать отклик на  $\delta$ -импульс, приложенный после этого момента).

Таким образом, передаточная функция  $\dot{K}(f)$  и импульсная переходная характеристика  $g(\tau)$  для линейных цепей с постоянными параметрами связаны преобразованиями Фурье. Обратное преобразование Фурье может быть записано в виде:

$$g(\tau) = \int_{-\infty}^{\infty} \dot{K}(f) e^{j2\pi f \tau} df. \quad (5.9)$$

Рассмотрим линейные преобразования случайных сигналов, ограничиваясь рамками корреляционной теории, когда требуется определить только корреляционную функцию или энергетический спектр случайного выходного сигнала и не требуется знать полностью определяющие его многомерные законы распределения.

В общем случае, когда входной случайный сигнал  $\xi(t)$  нестационарный, а линейная цепь имеет переменные параметры, корреляционная функция выходного случайного сигнала  $\eta(t)$  равна:

$$B_{\eta}(t_1, t_2) = m_1[\eta(t_1) \eta(t_2)] = m_1 \left[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(t_1, t'_1) \xi(t'_1) \times \right. \\ \left. \times g(t_2, t'_2) \xi(t'_2) dt'_1 dt'_2 \right] = \\ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(t_1, t'_1) g(t_2, t'_2) B_{\xi}(t'_2, t'_1) dt'_1 dt'_2.$$

Если входной сигнал  $\xi(t)$  стационарный, то

$$B_{\eta}(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(t_1, t'_1) g(t_2, t'_2) B_{\xi}(t'_2 - t'_1) dt'_1 dt'_2. \quad (5.10)$$

Из формулы (5.10) видно, что сигнал на выходе линейной системы с переменными параметрами нестационарен даже тогда, когда входной сигнал стационарен. Вводя новые переменные  $\tau_1 = t_1 - t'_1$  и  $\tau_2 = t_2 - t'_2$  из (5.10) для линейной системы с постоянными параметрами получаем:

$$B_{\eta}(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\tau_1) g(\tau_2) B_{\xi}(t_2 - \tau_2 - t_1 + \tau_1) d\tau_1 d\tau_2$$

и, вводя переменную  $\tau = t_2 - t_1$ , можно записать:

$$B_{\eta}(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\tau_1) g(\tau_2) B_{\xi}(\tau_1 - \tau_2 + \tau) d\tau_1 d\tau_2. \quad (5.11)$$

Таким образом, для линейной цепи с постоянными параметрами входному стационарному сигналу соответствует выходной стационарный сигнал, корреляционная функция которого  $B_{\eta}(\tau)$  полностью определяется корреляционной функцией  $B_{\xi}(\tau)$  входного стационарного сигнала и импульсной переходной характеристикой  $g(\tau)$  линейной цепи.

Условия стационарности входного сигнала, строго говоря, предполагают бесконечную его длительность. Для сигнала конечной длительности соотношение (5.11) выполняется лишь на участках стационарности, устанавливающихся после окончания переходных процессов в цепи.

Если стационарный случайный входной сигнал  $\xi(t)$  задан энергетическим спектром  $G_{\xi}(f)$ , то, используя связь между энергетическим спектром входного и выходного сигналов

$$G_{\eta} = |\dot{K}(f)|^2 G_{\xi}(f) = C^2(f) G_{\xi}(f) \quad (5.12)$$

и формулу (2.51), получаем:

$$B_{\eta}(\tau) = \int_{-\infty}^{\infty} G_{\eta}(f) \cos 2\pi f\tau df = \int_{-\infty}^{\infty} G_{\xi}(f) C^2(f) \cos 2\pi f\tau df. \quad (5.13)$$

Средняя мощность процесса на выходе линейной системы с постоянными параметрами на основании (5.11) и (5.13) равна

$$\sigma^2_{\eta} = B_{\eta}(0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} B_{\xi}(\tau_1 - \tau_2) g(\tau_1) g(\tau_2) d\tau_1 d\tau_2 = \int_{-\infty}^{\infty} G_{\xi}(f) C^2(f) df. \quad (5.14)$$

С помощью формул (5.10)–(5.14) полностью решается задача преобразования корреляционной функции или энергетического спектра стационарного случайного процесса при прохождении через линейную цепь.

### 5.3. Оптимальная линейная фильтрация по критерию минимума искажений полезного сигнала

Пусть для выделения полезного сигнала  $\xi(t)$  из входного сигнала

$$\eta = \xi(t) + n(t), \quad (5.15)$$

представляющего аддитивную смесь полезного сигнала с шумом  $n(t)$ , используется линейный фильтр, выходной сигнал которого  $\hat{\xi}(t)$  принимается в качестве оценки полезного сигнала. Выходной сигнал линейного фильтра может быть определен через импульсную переходную характеристику фильтра  $g(t, t')$ :

$$\hat{\xi}(t) = \int_{-\infty}^t g(t, t') \eta(t') dt'. \quad (5.16)$$

Задача оптимальной линейной фильтрации по критерию минимума искажения полезного сигнала сводится к выбору такой импульсной переходной характеристики фильтра  $g_{\text{опт}}(t, t')$ , которая минимизирует средний квадрат ошибки оценки

$$\sigma^2_{\xi}(t) = m_1 \{ [\xi(t) - \hat{\xi}(t)]^2 \}. \quad (5.17)$$

Эта задача решается методами вариационного исчисления. При этом искомая импульсная переходная характеристика оптимального фильтра  $g_{\text{опт}}(t, t')$  является решением следующего интегрального уравнения [1, 3]:

$$B_{\xi\eta}(t - \tau, t) = \int_0^{\infty} g_{\text{опт}}(t, t - \nu) B_{\eta}(t - \tau, t - \nu) d\nu, \quad (5.18)$$

а минимальная дисперсия ошибки оценки определяется формулой

$$\sigma^2_{\xi\min}(t) = B_{\eta}(t, t) - \int_0^{\infty} g_{\text{опт}}(t, t - \nu) B_{\xi\eta}(t - \nu, t) d\nu. \quad (5.19)$$

Таким образом, для решения задачи достаточно знать только корреляционные характеристики: взаимно-корреляционную функцию  $B_{\xi\eta}(t, t')$  полезного сигнала  $\xi(t)$  и входного сигнала  $\eta(t)$  и корреляционную функцию входного сигнала  $B_{\eta}(t, t')$ .

Если полезный сигнал и шум стационарны и стационарно связаны (именно для этого случая задача впервые была решена Н. Винером), то уравнение (5.18) и формула (5.19) принимают вид:

$$B_{\xi\eta}(\tau) = \int_0^{\infty} g_{\text{опт}}(\nu) B_{\eta}(\tau - \nu) d\nu, \quad (5.20)$$



$$\sigma_{\xi \min}^2 = B_{\eta}(0) - \int_0^{\infty} g_{\text{отт}}(\nu) B_{\xi\eta}(\nu) d\nu. \quad (5.21)$$

Если полезный сигнал и шум к тому же являются нормальными процессами, то линейный фильтр, удовлетворяющий интегральному уравнению (5.20), оказывается оптимальным и без ограничения класса оцениваемых фильтров линейными [3].

Решение интегральных уравнений (5.18) и (5.20) связано с довольно сложными преобразованиями [1, 3], а выражения для импульсной переходной характеристики и дисперсии оценки получаются весьма громоздкими и мало наглядными. Задача значительно упрощается, если, рассматривая случай стационарных и стационарно связанных полезного сигнала и шума, не накладывать на фильтр требования физической осуществимости, т.е. полагать  $g(\tau) \neq 0$  и при  $\tau < 0$ . Оптимальный линейный фильтр, полученный при такой идеализации импульсной переходной характеристики, носит название *идеального винеровского фильтра*. Для этого фильтра получаются очень простые и наглядные выражения для передаточной функции и дисперсии ошибки оценки полезной составляющей входного сигнала. И хотя эти характеристики соответствуют физически нереализуемому фильтру, они представляют большой интерес, определяя границу точности оценки сигнала, которая не может быть превышена физически реализуемыми линейными фильтрами (дополнительные ограничения на вид импульсной переходной характеристики физически реализуемых фильтров могут лишь сузить их возможности).

Определим передаточную функцию идеального винеровского фильтра и характеризующую его дисперсию ошибки оценки полезного сигнала, задавая полезный сигнал и шум, действующие на входе фильтра, значениями спектральной плотности мощности  $G_c(f)$  и  $G_m(f)$ . Для упрощения задачи полезный сигнал и шум будем считать независимыми. Будем искать передаточную функцию идеального винеровского фильтра в виде

$$\dot{K}(f) = C_{\phi}(f) e^{j\varphi_{\phi}(f)} = C_{\phi}(f) \cos \varphi_{\phi}(f) + j C_{\phi}(f) \sin \varphi_{\phi}(f), \quad (5.22)$$

где  $C_{\phi}(f)$  и  $\varphi_{\phi}(f)$  — амплитудно-частотная и фазочастотная характеристики фильтра, представляющие действительные функции.

Если бы шум отсутствовал, то для неискаженной передачи полезного сигнала нужно было бы, чтобы все составляющие спектра проходили без искажений, т.е. выполнялось условие  $\dot{K}(f) = 1$  для всех значений  $f$ , для которых  $G_c(f) \neq 0$ . Искажения сигнала на выходе фильтра при наличии шума состоят из двух составляющих. Составляющей от шума, прошедшего через фильтр. Дисперсия этой составляющей искажения сигнала в соответствии с (5.14) равна:

$$\sigma_m^2 = \int_{-\infty}^{\infty} |\dot{K}(f)|^2 G_m(f) df. \quad (5.23)$$

И составляющей, вызванной искажением полезного сигнала фильтром из-за отличия  $\dot{K}(f)$  от 1:

$$\sigma_c^2 = \int_{-\infty}^{\infty} |\dot{K}(f) - 1|^2 G_c(f) df. \quad (5.24)$$

В силу независимости этих искажений суммарная дисперсия ошибки оценки сигнала равна

$$\sigma_{\xi}^2 = 2 \int_0^{\infty} |\dot{K}(f)|^2 G_m(f) df + 2 \int_0^{\infty} |\dot{K}(f) - 1|^2 G_c(f) df. \quad (5.25)$$

С учетом (5.22) можем записать

$$\begin{aligned} \sigma_{\xi}^2 = 2 \int_0^{\infty} [C_{\phi}^2(f) G_m(f) + |C_{\phi}(f) \cos \varphi_{\phi}(f) + \\ + j C_{\phi}(f) \sin \varphi_{\phi}(f) - 1|^2 G_c(f)] df = 2 \int_0^{\infty} [C_{\phi}^2(f) G_m(f) + \\ + [C_{\phi}^2(f) - 2 C_{\phi}(f) \cos \varphi_{\phi}(f) + 1] G_c(f)] df. \end{aligned} \quad (5.26)$$

Поскольку подынтегральная функция в (5.26) положительна, интеграл принимает минимальное значение при максимальной абсолютной величине единственного отрицательного члена в подынтегральном выражении, т.е. при  $\cos \varphi_{\phi}(f) = 1$ . Это соответствует действительной передаточной функции фильтра

$$\dot{K}(f) = C_{\phi}(f), \quad (5.27)$$

обеспечивающей передачу отдельных спектральных составляющих входного сигнала без фазового сдвига.

При  $\cos \varphi_{\phi}(f) = 1$  формулу (5.26) можно преобразовать следующим образом:

$$\begin{aligned} \sigma_{\xi}^2 = 2 \int_0^{\infty} \left\{ [C_{\phi}(f) \sqrt{G_m(f) + G_c(f)}]^2 - 2 C_{\phi}(f) G_c(f) + \right. \\ \left. + \left[ \frac{G_c(f)}{\sqrt{G_m(f) + G_c(f)}} \right]^2 - \left[ \frac{G_c(f)}{\sqrt{G_m(f) + G_c(f)}} \right]^2 + G_c(f) \right\} df = \end{aligned}$$

$$= 2 \int_0^{\infty} \left\{ \left[ C_{\Phi}(f) \sqrt{G_m(f) + G_c(f)} - \frac{G_c(f)}{\sqrt{G_m(f) + G_c(f)}} \right]^2 + \frac{G_m(f) G_c(f)}{G_m(f) + G_c(f)} \right\} df. \quad (5.28)$$

Поскольку оба слагаемых под знаком интеграла в (5.28) неотрицательны, а второе слагаемое от  $C_{\Phi}(f)$  не зависит, оптимальное значение  $C_{\Phi}(f)$  должно обращать в нуль выражение в квадратных скобках в (5.28)

$$\dot{K}_{\text{опт}}(f) = C_{\text{опт}}(f) = \frac{G_c(f)}{G_c(f) + G_m(f)}. \quad (5.29)$$

Соответствующее значение дисперсии ошибки оценки полезного сигнала

$$\sigma_{\xi_{\min}}^2 = 2 \int_0^{\infty} \frac{G_m(f) G_c(f)}{G_m(f) + G_c(f)} df. \quad (5.30)$$

Поскольку  $G_c(f)/[G_c(f) + G_m(f)] < 1$ , то

$$\sigma_{\xi_{\min}}^2 < 2 \int_0^{\infty} G_m(f) df = \sigma_{\text{ш вх}}, \quad (5.31)$$

т.е. дисперсия ошибки оценки полезного сигнала при использовании идеального фильтра всегда меньше дисперсии ошибки, вызванной действием шума при отсутствии фильтрации.

В случае, если энергетические спектры полезного сигнала и шума нигде не перекрываются, то  $G_m(f) G_c(f) \equiv 0$  и искажения сигнала при идеальной фильтрации вообще отсутствуют. В общем же случае фильтрация тем эффективней, чем сильнее различаются энергетические спектры полезного сигнала и шума.

Следует отметить, что условие физической реализуемости фильтра вида  $g(\tau) = 0$  при  $\tau < 0$  не распространяется на системы, производящие неоперативную обработку записанных в памяти реализаций входного сигнала. При неоперативной обработке хранящихся в памяти достаточно длительных реализаций сигнала могут быть получены результаты, близкие к идеальной винеровской фильтрации, поскольку при оценке полезного сигнала могут использоваться достаточно протяженные участки реализаций входного сигнала как до, так и после момента времени, к которому относится оценка.

#### 5.4. Оптимальная линейная фильтрация по критерию максимума отношения сигнал-шум (согласованные фильтры)

Определим передаточную функцию линейного фильтра, обеспечивающего максимальное отношение мощности сигнала к мощности шума на выходе при входном сигнале, представляющем аддитивную смесь полезного сигнала заданной формы  $u(t)$  и стационарного шума  $n(t)$  с энергетическим спектром  $G_m(f)$ .

Запишем комплексный спектр полезного сигнала  $u(t)$  в виде

$$\dot{S}(f) = C_c(f) e^{j\varphi_c(f)}. \quad (5.32)$$

Обозначим передаточную функцию оптимального фильтра, обеспечивающего максимальное отношение сигнал/шум на выходе в некоторый момент времени  $t_0$ ,

$$\dot{K}_{\text{опт}}(f, t_0) = C_{\Phi \text{ опт}}(f) e^{j\varphi_{\Phi \text{ опт}}(f)}. \quad (5.33)$$

Значения отдельных спектральных составляющих полезного сигнала  $u(t)$  в момент  $t_0$  равны

$$\dot{S}(f) e^{j2\pi f t_0} = C_c(f) e^{j[2\pi f t_0 + \varphi_c(f)]}. \quad (5.34)$$

Из физических соображений очевидно, что оптимальная фазочастотная характеристика фильтра  $\varphi_{\Phi \text{ опт}}(f)$  соответствует условию сведения к нулю в момент  $t_0$  фаз всех спектральных составляющих полезного сигнала на выходе фильтра. В этом случае обеспечивается максимальное значение выходного сигнала, поскольку суммирование спектральных составляющих полезного сигнала на выходе фильтра в момент  $t_0$  водится к суммированию их амплитуд:

$$[u_{\text{вых max}}(t_0)] = \int_{-\infty}^{\infty} C_{\Phi \text{ опт}}(f) C_c(f) df. \quad (5.35)$$

Дисперсия шума на выходе фильтра не зависит от фазочастотной характеристики последнего: она полностью определяется энергетическим спектром выходного шума, не зависящим, как известно, от фазы спектральных составляющих. Поэтому принимаем:

$$\varphi_{\Phi \text{ опт}}(f) = -\varphi_c(f) - 2\pi f t_0. \quad (5.36)$$

Для определения амплитудно-частотной характеристики фильтра  $C_{\Phi \text{ опт}}(f)$ , максимизирующей отношение мощностей сигнала и шума на его выходе в момент  $t_0$ , оценим вклад в каждую из этих мощностей

элементарный участок спектра  $df$ . Вклад участка спектра  $df$  в величину  $u_{\text{вых max}}(t_0)$  согласно (5.35) равен

$$du_{\text{вых}} = C_{\text{ф опт}}(f) C_c(f) df$$

и соответствующий вклад в выходную мощность полезного сигнала

$$d\mathcal{P}_{\text{с вых}} = 2u_{\text{вых max}}(t_0) du_{\text{вых}} = 2u_{\text{вых max}}(t_0) C_{\text{ф опт}}(f) C_c(f) df. \quad (5.37)$$

Выходная мощность шума, приходящаяся на участок спектра  $df$ :

$$d\mathcal{P}_{\text{ш вых}} = G_{\text{ш}}(f) C^2_{\text{ф опт}}(f) df. \quad (5.38)$$

Оптимальная амплитудно-частотная характеристика фильтра, обеспечивающая максимум отношения сигнал/шум на его выходе, должна отвечать условию одинакового отношения элементарных мощностей  $d\mathcal{P}_{\text{с вых}}$  и  $d\mathcal{P}_{\text{ш вых}}$  для всех участков спектра. Если на каком-то участке спектра  $df$  отношение  $d\mathcal{P}_{\text{с вых}}/d\mathcal{P}_{\text{ш вых}}$  оказывается больше среднего, то для улучшения общего отношения сигнал/шум передаточную функцию на этой частоте было бы выгодно увеличивать до тех пор, пока отношение  $d\mathcal{P}_{\text{с вых}}/d\mathcal{P}_{\text{ш вых}}$  не снизилось бы до среднего уровня [из (5.37) и (5.38) видно, что с увеличением  $C_{\text{ф}}(f)$  это отношение уменьшается]. При нарушении этого условия в обратную сторону передаточную функцию на этом участке выгодно уменьшать до тех пор, пока отношение  $d\mathcal{P}_{\text{с вых}}/d\mathcal{P}_{\text{ш вых}}$  не поднимается до среднего уровня.

Условие постоянства по всему спектру отношения  $d\mathcal{P}_{\text{с вых}}/d\mathcal{P}_{\text{ш вых}}$  в соответствии с (5.37) и (5.38) имеет вид:

$$\frac{d\mathcal{P}_{\text{с вых}}}{d\mathcal{P}_{\text{ш вых}}} = \frac{2u_{\text{вых max}}(t_0) C_0(f)}{G_{\text{ш}}(f) C_{\text{ф опт}}(f)} = \text{const},$$

откуда следует, что

$$C_{\text{ф опт}}(f) = k C_c(f) / G_{\text{ш}}(f). \quad (5.39)$$

Опуская несущественный для определения передаточной функции фильтра постоянный масштабный коэффициент  $k$ , из (5.36) получаем:

$$\begin{aligned} \dot{K}_{\text{опт}}(f, t_0) &= \frac{C_c(f)}{G_{\text{ш}}(f)} \exp \{-j[\varphi_c(f) + 2\pi f t_0]\} = \\ &= \frac{\dot{S}_c^*(f)}{G_{\text{ш}}(f)} \exp(-j2\pi f t_0), \end{aligned} \quad (5.40)$$

где  $\dot{S}_c^*(f)$  — величина, комплексно-сопряженная спектру полезного сигнала  $\dot{S}_c(f)$ .

В случае, когда шум белый и  $G_{\text{ш}}(f) = N_0/2 = \text{const}$ , передаточная функция (5.40) принимает вид:

$$\dot{K}_{\text{опт}}(f, t_0) = \dot{S}_c^*(f) \exp(-j2\pi f t_0). \quad (5.41)$$

Таким образом, в случае белого шума передаточная функция линейного фильтра, максимизирующего отношение сигнал-шум на его выходе, должна быть согласована (комплексно сопряжена) со спектром полезного сигнала (сомножитель  $e^{-j2\pi f t_0}$  в (5.41) не влияет на форму выходного сигнала, а определяет лишь задержку  $t_0$  момента достижения максимума). Передаточную функцию (5.40), отвечающую случаю действия произвольного стационарного шума, можно считать согласованной со спектром полезной составляющей входного сигнала, подвергнутого предварительно операции "отбеливания" шума, отображаемой делителем  $G_{\text{ш}}(f)$ . Линейные фильтры, оптимальные по критерию максимума отношения сигнал/шум на выходе, благодаря этому свойству их передаточных функций, получили название согласованных фильтров.

Определим теперь импульсную передаточную характеристику согласованного фильтра  $g_{\text{опт}}(\tau; t_0)$ . В соответствии с (5.8)

$$\begin{aligned} g_{\text{опт}}(\tau; t_0) &= \int_{-\infty}^{\infty} \dot{K}_{\text{опт}}(f, t_0) e^{j2\pi f \tau} df = \\ &= \int_{-\infty}^{\infty} \dot{S}_c^*(f) e^{-j2\pi f t_0} e^{j2\pi f \tau} df = \int_{-\infty}^{\infty} \dot{S}_c(-f) e^{-j2\pi f(t_0 - \tau)} df = \\ &= \int_{-\infty}^{\infty} \dot{S}_c(f) e^{j2\pi f(t_0 - \tau)} df = u(t_0 - \tau). \end{aligned} \quad (5.42)$$

Для физически реализуемого фильтра полезный сигнал, подвергаемый фильтрации, должен задаваться на интервале  $[t_0 - T, t_0]$ , предшествующем моменту отсчета выходного сигнала  $t_0$ . Импульсная передаточная характеристика согласованного фильтра  $g_{\text{опт}}(\tau; t_0)$  получается, как следует из (5.42), зеркальным отображением сигнала  $u(t)$  относительно вертикальной оси, проходящей через точку  $t = t_0$ , если начало отсчета  $\tau$  совместить с этой точкой (рис. 5.2).

Определим максимальное отношение сигнал/шум на выходе согласованного фильтра  $u_{\text{вых}}^2(t_0)/\sigma_{\text{ш вых}}^2$  при действии белого шума. Используя передаточную функцию согласованного фильтра (5.41), получаем:

$$u_{\text{вых}}(t_0) = \int_{-\infty}^{\infty} \dot{K}_{\text{опт}}(f, t_0) \dot{S}_c(f) e^{j2\pi f t_0} df =$$

$$= \int_{-\infty}^{\infty} \dot{S}_c^*(f) e^{-j2\pi f t_0} \dot{S}_c(f) e^{j2\pi f t_0} df = \int_{-\infty}^{\infty} |\dot{S}_c(f)|^2 df = E_c, \quad (5.43)$$

где в соответствии с равенством Парсеваля  $E_c$  — энергия полезного сигнала на выходе фильтра.

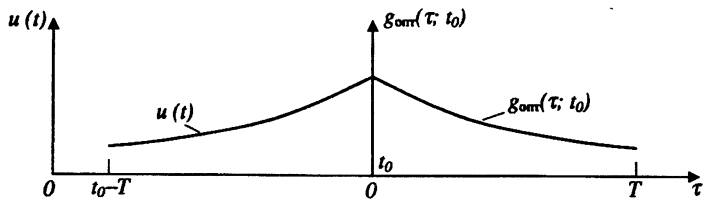


Рис. 5.2

Согласно (5.14) средняя мощность шума на выходе фильтра равна

$$\begin{aligned} \sigma_{\text{ш вых}}^2 &= \int_{-\infty}^{\infty} \frac{N_0}{2} |K_{\text{опт}}(f, t_0)|^2 df = \\ &= \frac{N_0}{2} \int_{-\infty}^{\infty} |\dot{S}_c(f)|^2 df = \frac{N_0}{2} E_c, \end{aligned} \quad (5.44)$$

где  $N_0/2$  — спектральная плотность мощности белого шума ( $N_0$  — спектральная плотность, соответствующая одностороннему спектру шума).

Из (5.43) и (5.44) следует, что

$$u_{\text{с вых}}^2(t_0)/\sigma_{\text{ш вых}}^2 = 2E_c/N_0.$$

Определим сигнальную составляющую на выходе согласованного фильтра  $u_{\text{с вых}}$ , соответствующую полезной составляющей входного сигнала  $u(t)$  при импульсной передаточной характеристике (5.41). В соответствии с (5.6)

$$u_{\text{с вых}}(t) = \int_0^{\infty} u(t-\tau) g_{\text{опт}}(\tau; t_0) d\tau = \int_0^{\infty} u(t-\tau) u(t_0-\tau) d\tau.$$

Введя новые переменные  $t_0 - \tau = t'$  и  $t_0 - t = t'$ , получим:

$$u_{\text{с вых}}(t) = \int_{-\infty}^{t_0} u(t') u(t' - t') dt' = \int_{-\infty}^{\infty} u(t') u(t' - t') dt', \quad (5.45)$$

поскольку предполагается, что  $u(t) = 0$  при  $t > t_0$ . Интеграл в правой части (5.45) представляет собой сигнальную составляющую корреляционного интеграла (4.17) для случая, когда оцениваемым параметром является временная задержка сигнала  $t'$ . Максимум полезной составляющей вы-

ходного сигнала согласованного фильтра, достигаемый при  $t = t_0$  совпадает с максимумом сигнальной составляющей корреляционного интеграла, достигаемым при  $t_0 - t = \tau' = 0$ . В обоих случаях отношение сигнальной и шумовой составляющих равно  $2E_c/N_0$ .

Таким образом, использование согласованного фильтра и корреляционного приемника дает эквивалентные результаты. Поэтому согласованный фильтр может использоваться для решения тех же задач, что и коррелятор, а именно: для оптимального обнаружения сигналов, оптимальной оценки их временной задержки, предельного сжатия по длительности широкополосных сигналов.

## 5.5. Примеры согласованных фильтров

В общем случае для сигнала  $u(t)$  произвольной формы техническая реализация согласованного фильтра представляет большие трудности. Даже для простых сигналов трудно обеспечить строгое выполнение условия согласования (5.41).

Рассмотрим в качестве примера простой сигнал в виде прямоугольного видеопульса (рис. 5.3):

$$u_{\text{вн}}(t) = \begin{cases} u_n & \text{при } -\tau_n/2 \leq t \leq \tau_n/2, \\ 0 & \text{при } |t| > \tau_n/2. \end{cases} \quad (5.46)$$

Комплексный спектр видеопульса

$$\dot{S}_{\text{вн}}(f) = \int_{-\tau_n/2}^{\tau_n/2} u_n e^{j2\pi f t} dt = \int_{-\tau_n/2}^{\tau_n/2} u_n \cos 2\pi f t dt = u_n \tau_n \frac{\sin \pi f \tau_n}{\pi f \tau_n} = u_n \tau_n \left| \frac{\sin \pi f \tau_n}{\pi f \tau_n} \right| e^{j\nu\pi}, \quad (5.47)$$

где  $\nu = [f\tau_n]_{\text{цел}}$  — целая часть  $f\tau_n$ .

Таким образом, для прямоугольного видеопульса

$$C_{\text{вн}}(f) = u_n \tau_n \frac{\sin \pi f \tau_n}{\pi f \tau_n}, \quad (5.48)$$

$$\varphi_{\text{вн}}(f) = \pi [f\tau_n]_{\text{цел}}. \quad (5.49)$$

Вид амплитудной и фазовой характеристик комплексного спектра прямоугольного видеопульса приведен на рис. 5.4 (штриховые линии). Амплитудная характеристика этого спектра  $C_{\text{вн}}(f)$  имеет многолепестковую форму. Функция  $C_{\text{вн}}(f)$  обращается в нуль в граничных точках между лепестками, отвечающих условию

$$f = \nu/\tau_n, \quad \nu = \pm 1, \pm 2, \dots, \quad (5.50)$$

и имеет максимумы в точках

$$f = 0 \text{ и } f = \frac{|v| + 1/2}{\tau_n} \text{ sign } v, \quad v = \pm 1, \pm 2, \pm 3, \dots, \quad (5.51)$$

где  $\text{sign } v$  — функция знака.

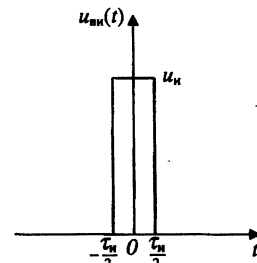


Рис. 5.3

Для центрального лепестка максимум  $C_{\text{вн}}(0)$  равен площади импульса  $u_n \tau_n$ , а для боковых лепестков имеет значения, затухающие по закону  $u_n \tau_n / (|v| + 1/2)\pi$ .

Фазовая характеристика комплексного спектра  $\varphi_{\text{вн}}(f)$  равна нулю для центрального лепестка и принимает значения  $v\pi$  ( $v = \pm 1, \pm 2, \dots$ ) для боковых лепестков. Сплошными линиями для  $f > 0$  (рис. 5.4) показана амплитудная характеристика одностороннего (действительного) спектра прямоугольного видеоимпульса, соответствующая его разложению по базисным функциям  $\cos 2\pi f t$ .

Амплитудно-частотная характеристика  $C_{\text{фонт}}(f)$  фильтра, согласованного со спектром прямоугольного видеоимпульса, должна воспроизводить функцию  $C_{\text{вн}}(f)$ , приведенную на рис. 5.4, а фазочастотная характеристика  $\varphi_{\text{фонт}}(f)$  должна отличаться лишь знаком от  $\varphi_{\text{вн}}(f)$ . Строгая реализация такой передаточной функции практически невозможна. Обычно ограничиваются согласованием (и то приближенным) передаточной функции фильтра с центральным лепестком спектра, в котором сосредоточена основная часть энергии сигнала. Такой фильтр должен иметь амплитудно-частотную характеристику, близкую к форме центрального лепестка спектра, а фазочастотную характеристику, близкую к нулю.

Определим теперь передаточную функцию линейного фильтра, согласованного со спектром

прямоугольного радиоимпульса. Сигнал  $u_{\text{рн}}(t)$ , представляющий прямоугольный радиоимпульс амплитуды  $u_n$  может быть представлен в комплексной форме:

$$u_{\text{рн}}(t) = \begin{cases} \frac{u_n}{2} [e^{j2\pi f_0 t} + e^{-j2\pi f_0 t}] & \text{при } -\tau_n/2 \leq t \leq \tau_n/2, \\ 0 & \text{при } |t| > \tau_n/2. \end{cases} \quad (5.52)$$

Комплексный спектр такого сигнала имеет вид

$$\begin{aligned} \dot{S}_{\text{рн}}(f) &= \int_{-\tau_n/2}^{\tau_n/2} u_{\text{рн}}(t) e^{-j2\pi f t} dt = \frac{u_n}{2} \left[ \int_{-\tau_n/2}^{\tau_n/2} e^{-j2\pi(f-f_0)t} dt + \right. \\ &\left. + \int_{-\tau_n/2}^{\tau_n/2} e^{-j2\pi(f+f_0)t} dt \right] = \frac{1}{2} \{ \dot{S}_{\text{вн}}[(f-f_0)] + \dot{S}_{\text{вн}}[(f+f_0)] \}. \end{aligned} \quad (5.53)$$

Комплексный спектр (5.53) состоит из двух симметричных относительно точки  $f = 0$  участков, получающихся смещением по оси частот спектра прямоугольного видеоимпульса в обе стороны на величину несущей частоты радиоимпульса  $f_0$  (штриховые линии на рис. 5.5). Односторонний гармонический спектр радиоимпульса (на рис. 5.5 он показан сплошными линиями) состоит из двух симметрично расположенных относительно несущей частоты  $f_0$  частей (боковых полос), повторяющихся по форме односторонний спектр видеоимпульса.

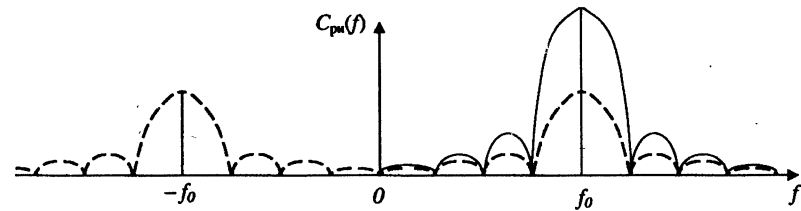


Рис. 5.5

Аналогично связаны между собой и передаточные функции линейных фильтров, согласованных с видеоимпульсом и радиоимпульсом. Для радиоимпульса обычно также ограничиваются согласованием передаточной функции фильтра с центральным лепестком спектра, занимающим полосу  $\pm 1/\tau_n$  около несущей частоты  $f_0$  (ширина полосы фильтра для радиоимпульса вдвое больше, чем для видеоимпульса той же длительности).

В качестве третьего примера рассмотрим согласованный фильтр для широкополосного сигнала типа ФМПС, уже приводившегося в § 4.6. В соответствии с (4.6) этот сигнал может быть записан в виде:

$$u_{\text{рн}}(t) = \begin{cases} \cos \left[ 2\pi f_0 t + (d_i - 1) \frac{\pi}{2} \right] & \text{при } 0 \leq t \leq T, \\ 0 & \text{при } t < 0 \text{ и } t > T, \end{cases} \quad (5.54)$$

где  $d_i$  — символы  $m$ -последовательности, принимающие значения  $\pm 1$  или  $-1$ ;  $T$  — длительность ФМПС.

Комплексный спектр сигнала (5.54) можно представить в виде:

$$\begin{aligned} \dot{S}_{\text{рн}}(f) &= \frac{1}{2} \int_0^T \exp \left\{ j \left[ 2\pi f_0 t + (d_i(t) - 1) \frac{\pi}{2} \right] \right\} \exp(-j2\pi f t) dt + \\ &+ \frac{1}{2} \int_0^T \exp \left\{ -j \left[ 2\pi f_0 t + (d_i(t) - 1) \frac{\pi}{2} \right] \right\} \exp(-j2\pi f t) dt = \\ &= \frac{1}{2} \sum_{i=0}^{N-1} \left\{ \exp \left\{ -j \left[ -(d_i - 1) \frac{\pi}{2} + 2\pi(f - f_0)t_i \right] \right\} \times \right. \\ &\times \int_{t_i}^{t_{i+1}} \exp[-j2\pi(f - f_0)(t - t_i)] dt + \exp \left\{ -j \left[ (d_i - 1) \frac{\pi}{2} + \right. \right. \\ &\left. \left. + 2\pi(f + f_0)t_i \right] \right\} \int_{t_i}^{t_{i+1}} \exp[-j2\pi(f + f_0)(t - t_0)] dt \left. \right\}, \end{aligned} \quad (5.55)$$

где  $t_i = i\tau$ ;  $\tau$  — шаг манипуляции фазы в ФМПС.

Спектр (5.55) представлен в виде суммы спектров, примыкающих друг к другу  $N$  элементарных радиоимпульсов длительностью  $\tau_0$  на которые может быть разбит ФМПС. Полагая для простоты шаг  $\tau_0$  фазовой манипуляции кратным периоду несущей частоты и  $\exp(-j2\pi f_0 t_i) = 1$  и вводя новую переменную интегрирования  $t - t_i = t'$ , получаем

$$\dot{S}_{\text{рн}}(f) = \frac{1}{2} \sum_{i=0}^{N-1} \exp \left\{ -j \left[ 2\pi f t_i - (d_i - 1) \frac{\pi}{2} \right] \right\} \times$$

$$\times \int_0^{\tau_0} \exp[-j2\pi(f-f_0)t'] dt' + \frac{1}{2} \sum_{i=0}^{N-1} \exp\left\{-j\left[2\pi f t_i + (d_i - 1) \frac{\pi}{2}\right]\right\} \int_0^{\tau_0} \exp[-j2\pi(f+f_0)t'] dt'. \quad (5.56)$$

Передаточная функция линейного фильтра, согласованного с ФМПС, имеет вид:

$$\begin{aligned} K_{\text{лн}}(f) &= S_{\text{лн}}^*(f) \exp(-j2\pi f t_0) = \\ &= \frac{1}{2} \sum_{i=0}^{N-1} \exp\left\{j\left[2\pi f(t_0 - t_i) - (d_i - 1) \frac{\pi}{2}\right]\right\} \int_0^{\tau_0} \exp[j2\pi(f-f_0)t'] dt' + \\ &+ \frac{1}{2} \sum_{i=0}^{N-1} \exp\left\{j\left[2\pi f(t_0 - t_i) + (d_i - 1) \frac{\pi}{2}\right]\right\} \int_0^{\tau_0} \exp[j2\pi(f+f_0)t'] dt'. \end{aligned} \quad (5.57)$$

Согласованный фильтр с передаточной функцией вида (5.57) можно представить в виде двух последовательно включенных линейных цепей. Первая из них осуществляет задержку элементарного радиоимпульса, поступающего на вход фильтра в момент  $t_i$ , на время  $t_0 - t_i$  и снимает манипуляцию фазы его высокочастотного заполнения, имевшую место при формировании ФМПС. В момент  $t_0$  на выходе этой цепи все элементарные импульсы оказываются совмещенными, а их высокочастотное заполнение суммируется по амплитуде, в результате чего формируется суммарный радиоимпульс длительностью  $\tau_0$ . Вторая цепь представляет линейный фильтр, согласованный с этим радиоимпульсом.

Такой двухзвенный фильтр может быть реализован с использованием широкополосной линии задержки с малым затуханием и общим временем задержки  $t_0$ , равным периоду  $m$ -последовательности  $T$ , и последовательно соединенного с ней фильтра, согласованного с радиоимпульсом длительностью  $\tau_0$  (рис. 5.6). При этом требуемая задержка элементарных радиоимпульсов, составляющих ФМПС, обеспечивается выбором отводов с линии задержки с шагом, равным  $\tau_0$ , а компенсация фазовой манипуляции — соответствующим выбором полярности трансформаторов, используемых для съема сигналов с отводов линии на сумматор.

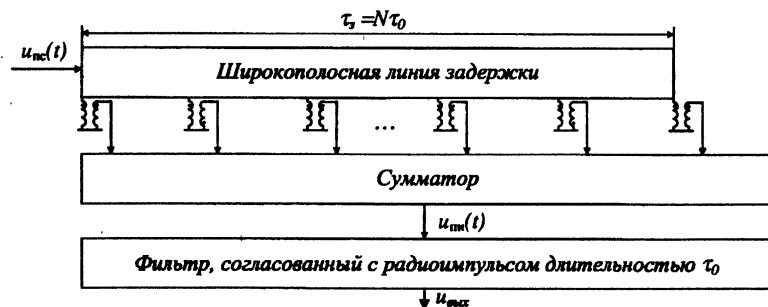


Рис. 5.6

Рассмотрим в качестве исходного момент, когда символ  $d_1$  достиг конца широкополосной линии задержки, а символ  $d_N$  поступил на ее вход. Пусть именно для этого положения сигнала полярности трансформаторов согласованы с законом фазовой манипуляции ФМПС. Тогда в этот момент времени элементарные сигналы, снимаемые с отводов линии задержки, будут суммироваться на выходной нагрузке по амплитуде. Если длительность сигнала составляет несколько периодов  $m$ -последовательности, то такая картина будет повторяться через каждый ее период. При смещении же сигнала проходящего через линию задержки, относительно исходного положения на интервал, некратный  $N\tau_0$ , в соответствии со свойствами  $m$ -последовательности (см. § 5.4) для  $(N-1)/2$  отводов сохранится та же полярность сигнала, что и при исходном положении, а для других  $(N+1)/2$  отводов полярность изменится на  $\pi$ . В результате компенсации элементарных радиоимпульсов с противофазными сигналами амплитуда сигнала на выходе линии задержки уменьшится в  $N$  раз, а его фаза изменится на  $\pi$ . Таким образом, выходной сигнал повторяет форму корреляционной функции ФМПС, приведенную на рис. 4.9, что подтверждает эквивалентность согласованного фильтра коррелятору. Сформированный на выходе линии задержки радиоимпульс поступает на согласованный с ним фильтр, обеспечивающий максимальное значение отношения сигнал/шум на его выходе.

Не следует забывать, что вывод об эквивалентности применения корреляционного приемника и согласованного фильтра справедлив только по отношению к задаче оптимального выделения сигнала на фоне белого шума. При реализации же систем, решающих определенные технические задачи, эти два варианта отнюдь не эквивалентны и выбор может быть сделан в каждом отдельном случае с учетом конкретных требований к системе, к числу которых относятся требования к времени поиска сигнала, к защищенности системы от организованных или внутрисистемных помех (возможности изменения формы сигнала), сложность технической реализации устройства при выбранной форме сигнала.

По времени поиска выигрыш дает согласованный фильтр, который всегда готов к приему сигнала данной формы. Эталонный сигнал отображен в его передаточной функции. Выходной сигнал согласованного фильтра воспроизводит непрерывный ряд значений корреляционного интеграла, отвечающих отрезкам входного сигнала длительностью  $T$ , предшествующим моменту отсчета, а момент поступления в фильтр участка сигнала, совпадающего по форме с эталоном, фиксируется максимальным всплеском выходного сигнала. В корреляционных одноканальных приемниках требуется определенное время на поиск задержки местного сигнала, обеспечивающей максимум его корреляции с принимаемым сигналом. Однако при этом следует сделать две оговорки: 1) это преимущество согласованных фильтров не относится к поиску сигнала по частоте; при смещении средней частоты спектра согласованный фильтр, рассчитанный на заданный спектр сигнала, требует точно такого же поиска по частоте, как и корреляционный приемник; 2) если для повышения помехозащищенности используется временное стробирование выходного сигнала согласованного фильтра, требуется и поиск по времени задержки, только в этом случае вместо задержки местного сигнала осуществляется задержка строга, представляющая более простую задачу.

По помехозащищенности от организованных помех в случае, если на выходе согласованного фильтра вводится временное стробирование, обе схемы практически равноценны. При этом после захвата сигнала в обоих случаях для эффективного воздействия на канал помеха должна не только воспроизводить форму, но и соответствовать полезному сигналу по временной задержке. При вхождении же в связь в обоих случаях может быть захвачена помеха, воспроизводящая форму сигнала, но не его задержку.

По возможности технической реализации иногда проще согласованный фильтр, иногда — корреляционный приемник. Так, например, при использовании ФМПС с малыми значениями базы сигнала более простым может оказаться согласованный фильтр; при больших же значениях базы, когда следует использовать широкополосные линии задержки с большим временем задержки и большим числом отводов, проще корреляционный приемник.

Наконец, в тех случаях, когда требуется гибкость системы (в смысле возможности изменения формы сигнала в процессе ее эксплуатации), предпочтение отдается корреляционным приемникам, особенно при применении ФМПС.

Следует отметить, что при сравнении согласованных фильтров и корреляционных приемников имелись в виду аналоговые фильтры. При использовании дискретной техники с применением цифровых ЭВМ для обработки информации (цифровые фильтры) различие между согласованными фильтрами и корреляционными приемниками стирается.

## II ОСНОВЫ ТЕОРИИ ИНФОРМАЦИИ И ИДЕАЛЬНОЕ КОДИРОВАНИЕ

### Глава 6 КОЛИЧЕСТВЕННОЕ ОПРЕДЕЛЕНИЕ ИНФОРМАЦИИ И ИНФОРМАЦИОННЫЕ ХАРАКТЕРИСТИКИ ИСТОЧНИКОВ СООБЩЕНИЙ

#### 6.1. Количественное определение информации, содержащейся в сообщении

Задачей любого канала связи является передача на расстояние *информации*, т.е. совокупности сведений о каком-либо событии, процессе, состоянии и т.п. Совокупность сведений, подлежащих передаче, представленная в виде соответствующей последовательности символов, называется *сообщением*. Объект, к которому относятся эти сведения, будем называть *источником сообщения*.

Количественная мера информации, содержащейся в сообщении, должна характеризовать степень уменьшения неопределенности наших представлений о состоянии источника сообщений после получения данного сообщения. Такая мера была введена основоположником современной теории информации известным американским ученым К. Шенноном [24].

Рассмотрим сначала вопрос о количественной мере информации применительно к источникам дискретных сообщений. Пусть источник дискретных сообщений выдает последовательность элементарных сообщений  $\{a_i\}$ , каждое из которых соответствует одному из возможных текущих его состояний. Совокупность элементарных сообщений  $a_1, a_2, \dots, a_m$ , отвечающих возможным состояниям источника, называется *алфавитом*, а число  $m$  различных элементарных сообщений — *объемом алфавита* источника. При передаче текста — это обычный алфавит данного языка; при передаче команд — перечень возможных команд; при передаче цифровых кодов — символы принятой системы счисления; при передаче сообщения о состоянии системы — перечень возможных состояний системы или ее составных частей и т.д. В дальнейшем для краткости элементарные сообщения будем называть символами. В случае, когда для передачи отдельных символов используется сложная кодовая последовательность, элементы

последней, в отличие от символов алфавита, будем называть элементарными символами.

Рассмотрим сначала случай достоверного приема сообщения, когда после получения символа  $a_i$  состояние источника сообщений на момент передачи этого символа определяется однозначно. Информация, доставляемая символом  $a_i$  будет тем больше, чем неожиданнее для получателя оказалось соответствующее состояние источника сообщений. Если по априорной информации (включая и информацию, доставленную предшествующими символами) получателю было достоверно известно, что в момент передачи данного символа источник находился в состоянии  $a_i$ , т.е. априорная вероятность этого состояния равна единице, то получение символа  $a_i$  не даст получателю никакой дополнительной информации. Если эта вероятность отлична от единицы, но близка к ней, то получение символа  $a_i$  несущественно уточняет наши представления о состоянии источника сообщений и вносит мало информации. Чем меньше априорная вероятность состояния  $a_i$  тем большую информацию вносит получение символа  $a_i$ . Поэтому можно считать, что количественная мера информации, которую несет символ  $a_i$ , должна быть функцией априорной вероятности  $P_i$  того, что источник в момент его передачи находился в состоянии  $a_i$ :

$$\mathcal{I}(a_i) = \varphi(P_i), \quad (6.1)$$

где  $P_i = P(a_i|a_j a_k \dots)$  — условная вероятность получения символа  $a_i$  после поступления конкретной совокупности символов  $a_j, a_k, \dots$ . Для упрощения записи мы используем обозначение  $P_i$ , однако следует помнить, что это условная вероятность, зависящая не только от общих свойств источника сообщений, но и от конкретных значений переданных ранее символов.

Определим вид функции  $\varphi(P_i)$ , потребовав, чтобы количество информации, помимо отмеченного условия

$$\mathcal{I}(a_i) = \varphi(1) = 0 \text{ при } P_i = 1, \quad (6.2)$$

удовлетворяло бы еще условию аддитивности, отвечающему нашим интуитивным представлениям об этой мере [17]. Последнее означает, что при последовательной передаче символов  $a_i$  и  $a_j$  количество информации, вносимое этой парой символов, равно сумме количеств информации, вносимого символом  $a_i$  и добавляемого символом  $a_j$ :

$$\varphi[P(a_i, a_j)] = \varphi(P_i P_j) = \varphi(P_i) + \varphi(P_j), \quad (6.3)$$

где  $P(a_i, a_j)$  — условная вероятность поступления пары символов  $a_i$  после приема конкретной последовательности предшествовавших им символов:  $P_j = P(a_j|a_i, \dots)$  — условная вероятность поступления символа  $a_j$ , если перед этим был принят символ  $a_i$  и та же последовательность предшествовавших ему символов.

Продифференцировав (6.3) по  $P_i$  имеем

$$P_j \varphi'(P_i P_j) = \varphi'(P_i),$$

а умножив обе части на  $P_i$  получим

$$P_i P_j \varphi'(P_i P_j) = P_i \varphi'(P_i). \quad (6.4)$$

Равенство (6.4) должно выполняться при произвольных значениях  $P_i$  и  $P_i P_j$ , что означает инвариантность обеих его частей к значениям аргумента функции  $\varphi$ :

$$P_i P_j \varphi'(P_i P_j) = P_i \varphi'(P_i) = \text{const} = k, \quad (6.5)$$

откуда

$$\varphi(P_i) = k \ln P_i + C. \quad (6.6)$$

Из условия  $\varphi(1) = 0$  получим, что произвольная постоянная  $C = 0$  и, следовательно,

$$\mathcal{I}(a_i) = k \ln P_i. \quad (6.7)$$

Масштабный коэффициент  $k$  зависит от выбора единицы измерения количества информации.

Примем за единицу количество информации, содержащееся в сообщении о том, что произошло одно из двух равновероятных независимых событий (такая единица количества информации называется двоичной). Тогда получим

$$k \ln \frac{1}{2} = 1; \quad k = \ln^{-1} \frac{1}{2} = -\ln^{-1} 2$$

и соответственно

$$\mathcal{I}(a_i) = -\ln P_i / \ln 2 = -\log_2 P_i. \quad (6.8)$$

Свойство аддитивности позволяет на основании формулы (6.8) для информации, переносимой одним символом, определить количество информации, содержащейся в любом сколь угодно длинном сообщении:



$$\mathcal{I}(a_{i_1}, a_{i_2}, \dots, a_{i_q}) = - \sum_{v=1}^q \log_2 P_{i_v}, \quad (6.9)$$

где  $P_{i_v}$  — условная вероятность того, что  $v$ -м по порядку в сообщении (при данной совокупности предшествующих символов) будет символ  $a_{i_v}$ . В принятых здесь обозначениях индекс  $v$  определяет положение (порядковый номер) символа  $a_{i_v}$  в сообщении; индекс  $i_v$  принимающий одно из значений в пределах от 1 до  $m$ , определяет конкретный выбор этого символа.

## 6.2. Информационные характеристики источников дискретных сообщений

Формулы (6.8) и (6.9) определяют количество информации, переносимой при достоверном приеме отдельным символом или группой символов в конкретной реализации сообщения. Количество информации выступает здесь как случайная величина, характеризующая конкретные реализации сообщений. Нас же интересуют общие статистические свойства информации, поступающей от данного источника сообщений.

В качестве общей количественной характеристики информации, поступающей от конкретного источника сообщения  $A$ , в задачах связи удобно использовать простейшую числовую вероятностную характеристику — математическое ожидание случайной величины  $\mathcal{I}(a_i)$ , представляющую среднее количество информации  $\mathcal{I}(A)$ , приходящейся на один символ данного источника сообщений (последний полагается стационарным). Это оправдывается тем, что в задачах связи, как правило, оптимизируются именно среднее количество и средняя достоверность передачи информации по каналу за достаточно длительные интервалы времени. Случайные отклонения от среднего значения количества информации, переданного за конкретный интервал времени, не учитываемые этой характеристикой, могут иметь значение лишь в радиотелиографиях с особо жесткими допусками на запаздывание передачи информации. Только в этих специальных случаях, которые мы не будем рассматривать, величина  $\mathcal{I}(A)$  может оказаться недостаточно полной статистической характеристикой сообщений, поступающих от источника  $A$ .

В общем случае, когда условная вероятность  $P_{i_q}$  появления символа  $a_{i_q}$  зависит от состава предшествующих  $(q-1)$  символов, осреднение при определении математического ожидания должно производиться по всем возможным сочетаниям из  $q$  символов:

$$\mathcal{I}(A) = - \sum_{i_1=1}^m \dots \sum_{i_q=1}^m P(a_{i_1}, a_{i_2}, \dots, a_{i_q}) \log_2 P_{i_q}, \quad (6.10)$$

где  $P(a_{i_1}, a_{i_2}, \dots, a_{i_q})$  — совместная вероятность появления последовательности символов  $a_{i_1}, a_{i_2}, \dots, a_{i_q}$  — условная вероятность появления символа  $a_{i_q}$  после получения символов  $a_{i_1}, \dots, a_{i_{q-1}}$ . В частности, если вероятность появления очередного символа  $a_i$  зависит только от вероятности появления предыдущего символа  $a_j$ , формула (6.10) принимает вид:

$$\begin{aligned} \mathcal{I}(A) &= - \sum_{i=1}^m \sum_{j=1}^m P(a_i, a_j) \log_2 P(a_i|a_j) = \\ &= - \sum_{i=1}^m \sum_{j=1}^m P(a_j) P(a_i|a_j) \log_2 P(a_i|a_j). \end{aligned} \quad (6.11)$$

В простейшем случае, когда символы независимы, получаем

$$\mathcal{I}(A) = \sum_{i=1}^m P_i \mathcal{I}(a_i) = - \sum_{i=1}^m P_i \log_2 P_i. \quad (6.12)$$

Величины, стоящие в правых частях выражений (6.10)–(6.12), зависят только от вероятностных характеристик источника сообщений. Они определяют среднюю неопределенность очередного состояния источника и носят название *информационной энтропии источника сообщений* на один символ или, сокращенно, *энтропии источника сообщений*. Будем обозначать ее  $H(A)$ .

До сих пор мы рассматривали случай достоверной передачи символов сообщения. В этом случае апостериорная (после получения символа) неопределенность состояния источника сообщений отсутствует и количество информации, доставляемое в среднем одним символом, равно информационной энтропии источника сообщений:

$$\mathcal{I}(A) = H(A). \quad (6.13)$$

В дальнейшем, во избежание громоздкости записи условно будем использовать простейшую форму представления энтропии, соответствующую независимым символам источника сообщений:

$$H(A) = - \sum_{i=1}^m P_i \log_2 P_i,$$

имея в виду, что более общий случай отличается лишь формой осреднения.

Если в канале передачи информации возможны искажения символов (канал с помехами), то условие (6.13) уже не соблюдается. В этом случае текущее состояние источника сообщения после приема очередного символа  $a'_j$

определяется не достоверно, а характеризуется некоторым законом распределения  $P(a_i|a')$ . Соответственно апостериорная (после приема символа  $a'$ ) энтропия источника сообщений отлична от нуля и количество информации, приходящееся в среднем на один принятый символ, характеризуемое достигнутым уменьшением энтропии источника,

$$\mathcal{I}(A) = H(A) - H(A|A'), \quad (6.14)$$

где  $H(A|A')$  — апостериорная энтропия источника  $A$  после приема очередного символа искаженного помехами сообщения  $A'$ . Характер зависимости величины  $H(A|A')$  от вероятностей искажения символов будет рассмотрен в § 7.2.

Покажем, что максимум информационной энтропии источника сообщений достигается в случае, когда  $P_i = 1/m$ , т.е. когда источник характеризуется равной вероятностью и независимостью выбора символов  $a_i$  алфавита. Найдем условный экстремум величины

$$H(A) = - \sum_{i=1}^m P_i \log_2 P_i \quad (6.15)$$

при условии

$$\sum_{i=1}^m P_i = 1. \quad (6.16)$$

Для этого воспользуемся методом неопределенного множителя Лагранжа [38], позволяющим свести задачу определения условного экстремума функции  $f(x_1, x_2, \dots, x_m) = 0$  при условии  $\varphi(x_1, x_2, \dots, x_m) = 0$  к определению безусловного экстремума функции

$$F(x_1, x_2, \dots, x_m, \lambda) = f(x_1, x_2, \dots, x_m) + \lambda \varphi(x_1, x_2, \dots, x_m). \quad (6.17)$$

Применяя этот метод для определения максимума выражения (6.15) при условии (6.16), сводим задачу к определению максимума функции вида:

$$F(P_1, P_2, \dots, P_m, \lambda) = - \sum_{i=1}^m P_i \log_2 P_i + \lambda \left[ \sum_{i=1}^m P_i - 1 \right],$$

т.е. к выполнению условий

$$\frac{\partial}{\partial P_i} F(P_1, P_2, \dots, P_m, \lambda) = -\log_2 P_i - P_i \frac{1}{P_i \ln 2} + \lambda = 0$$

или

$$\log_2 P_i = \lambda - \ln^{-1} 2 = \text{const},$$

откуда  $P_i = \text{const} = 1/m$ . Подставив это значение в (6.15), получим:

$$H_{\max}(A) = -m \frac{1}{m} \log_2 \frac{1}{m} = \log_2 m. \quad (6.18)$$

Заметим, что одинаковая вероятность поступления любого из символов алфавита при передаче сообщения предполагает отсутствие корреляции между ними, ибо, по определению  $P_i$  представляет условную вероятность, учитывающую конкретные значения предшествующих символов. Естественно, что при наличии корреляции между символами значение  $P_i$  зависит от конкретных значений предшествовавших символов сообщения и условие  $P_i = \text{const}$  не может выполняться. Корреляция между символами всегда приводит к снижению среднего количества информации, доставляемой одним символом, ибо получение предшествующих символов при наличии корреляции уменьшает неопределенность выбора очередного символа. Так, например, поступление гласной буквы при передаче текста исключает вероятность поступления таких букв, как *ь, ъ, ы*, а для остальных букв усиливает неравномерность вероятностей поступления, повышая ее для одних и понижая для других букв. Все это приводит к уменьшению средней информации, доставляемой очередным символом.

Таким образом, наиболее экономичным является алфавит, использующий некоррелированные равновероятные символы. Любой другой алфавит при том же его объеме  $m$  потребует большего числа символов на передачу того же объема информации. Если источник сообщения с объемом алфавита  $m$  характеризуется энтропией  $H(A)$ , то среднее количество информации, содержащейся в  $n$  символах, поступивших от этого источника, равно  $nH(A)$ . Минимальное же число символов, необходимое для передачи того же объема информации при использовании алфавита с равновероятными некоррелированными символами, равно

$$n_{\min} = nH(A)/H_{\max}(A) = nH(A)/\log_2 m.$$

Избыточность числа символов, используемых данным источником сообщений для передачи некоторого количества информации, относительно минимально необходимого их числа, соответствующего использованию равновероятных независимых символов, характеризуется коэффициентом

$$\rho_n = (n - n_{\min})/n = 1 - H(A)/\log_2 m, \quad (6.19)$$

получившим название *избыточности* источника.

Было бы неправильным всегда рассматривать избыточность как признак несовершенства источника сообщений. Так, при передаче текста, когда свойства источника сообщений определяются словарем данного языка,

избыточность совершенно необходима для обеспечения удобства заучивания и произношения слов, разборчивости речи, мелодичности языка и т.п. Невозможно представить себе язык без избыточности, использующий слова из всех возможных сочетаний букв. Например, для русского алфавита, принимая  $m = 32$ , получаем максимально возможную энтропию

$$H_{\max}(A) = \log_2 32 = 5 \text{ дв. ед./симв.}$$

Это средняя информация на символ для "абсолютно хаотического текста", когда любое сочетание букв включено в словарь языка. В тексте же, соответствующем нормальной русской речи,  $H(A) = 1,5$  дв. ед./симв. [17], т.е. в три раза меньше, что соответствует избыточности источника сообщений (текста)  $\rho_{\text{из}} = 1 - 1,5/5 = 0,7$ . В общем случае передачи информации по каналам связи избыточность является необходимой платой за достоверность передачи информации в условиях наличия помех.

Отметим еще одно важное обстоятельство, вытекающее из формулы (6.12), касающееся среднего вклада маловероятных символов в передачу информации. Несмотря на то, что в каждом конкретном случае поступления таких символов имеет место наибольшее приращение информации, в среднем эти символы, именно благодаря малой вероятности их появления, вносят малый вклад в передачу информации, а при  $P_i \rightarrow 0$  этот вклад стремится к нулю. Действительно, применяя для раскрытия неопределенности правило Лопиталья, получаем:

$$\lim_{P_i \rightarrow 0} P_i \log_2 P_i = \lim_{P_i \rightarrow 0} \frac{d \log_2 P_i / d P_i}{d(1/P_i) / d P_i} = 0.$$

Это позволяет при статистической характеристике передаваемой информации исключить из рассмотрения те сообщения, суммарная вероятность которых стремится к нулю, чем мы и будем пользоваться далее при доказательстве теорем Шеннона.

Следует отметить, что введенная количественная мера информации совершенно не учитывает полезность, ценность или важность сообщений. Так, сообщение об оценке, полученной в школе учеником, неопределенность априорной информации о которой может быть достаточно большой, при использовании этой меры может содержать большее количество информации, чем сообщение о состоянии здоровья ребенка. Именно абстрагирование от качественных характеристик сообщения позволило создать строгий и изящный математический аппарат для характеристики информации. Однако в некоторых областях применения (в частности, в теории массового обслуживания), где учет полезности и важности информации существен, эта особенность меры информации, введенной Шенноном, создает определенные затруднения в использовании аппарата теории ин-

формации. Этот недостаток обычно исправляют разделением сообщений на категории с введением соответствующих приоритетов или весовых коэффициентов, однако при этом теряется общность и математическая строгость теории (в силу субъективности соответствующих показателей).

В задачах связи абстрагирование от качественного содержания информации представляется в значительной мере оправданным тем, что и канал связи индифферентен к содержанию передаваемой информации. Поэтому именно в области связи применение теории информации в классическом ее виде оказалось наиболее плодотворным. Важность и ценность информации может учитываться при организации работы системы связи как системы массового обслуживания введением категорий и приоритетов сообщений, а в некоторых случаях — и разных требований к надежности их передачи (и соответственно разных режимов работы каналов). Однако при этом введенные информационные характеристики полностью сохраняют свой смысл, по крайней мере, в рамках сообщений одной категории, и аппарат теории информации поэтому полностью применим.

### 6.3. Энтропия источников непрерывных сообщений

Источники непрерывных сообщений характеризуются тем, что в каждый момент времени  $t$  сообщение  $x(t)$  может принимать бесконечное множество значений с бесконечно малой вероятностью каждого из них, и если бы сообщение могло передаваться абсолютно точно без искажений, оно несло бы бесконечное количество информации. Однако на практике при передаче информации всегда имеют место искажения, и количество информации, содержащееся в принятом непрерывном сообщении, определяется разностью значений энтропии источника сообщений до и после получения информации. Эта разность, в отличие от абсолютного значения энтропии непрерывного источника, оказывается конечной величиной.

Пусть возможные значения сигнала  $x(t)$  в момент  $t$  характеризуются плотностью вероятности  $w(t)$ . Осуществим квантование возможных значений  $x(t)$  с шагом  $\Delta x$  (обратный переход к непрерывному распределению реализуется при  $\Delta x \rightarrow 0$ ). Вероятность того, что значение  $x$  лежит в пределах  $i$ -го шага квантования  $P_i = w(x) \Delta x$ . Тогда в соответствии с (6.14) информация на один отсчет такого квантованного сигнала равна (отсчеты полагаем некоррелированными):

$$H_{\Delta x}(x) = - \sum_{(i)} [w(x_i) \Delta x] \log_2 [w(x_i) \Delta x] =$$

$$\begin{aligned}
 &= - \sum_i [w(x_i) \log_2 w(x_i)] \Delta x - \log_2 \Delta x \sum_i w(x_i) \Delta x = \\
 &= \sum_i w(x_i) \log_2 w(x_i) \Delta x - \log_2 \Delta x,
 \end{aligned}$$

так как

$$\sum_i w(x_i) \Delta x = 1.$$

Переходя к пределу при  $\Delta x \rightarrow 0$ , получаем, что энтропия на один отсчет непрерывного источника сообщений равна

$$\begin{aligned}
 H(x) &= \lim_{\Delta x \rightarrow 0} H \Delta x(x) = \\
 &= - \int_{-\infty}^{\infty} w(x) \log_2 w(x) dx - \lim_{\Delta x \rightarrow 0} \log_2 \Delta x.
 \end{aligned} \quad (6.20)$$

Первый член правой части (6.20)

$$h(x) = - \int_{-\infty}^{\infty} w(x) \log_2 w(x) dx \quad (6.21)$$

называется *дифференциальной энтропией* непрерывного распределения (источника непрерывного сообщения). Величина  $h(x)$  имеет конечное значение. Она не зависит от шага квантования  $\Delta x$ , а зависит только от закона распределения непрерывной величины  $x$ . Второй член правой части (6.20), напротив, не зависит от характеристик источника сообщений, а определяется только шагом квантования  $\Delta x$  уровня сигнала, и именно этот член обращает значение  $H(x)$  для непрерывного источника в бесконечность.

Поскольку абсолютно точный отсчет значений непрерывного сигнала  $x(t)$  принципиально невозможен, то возможные состояния источника сообщений после получения отсчета  $x'$  будут определяться некоторым непрерывным законом распределения  $w(x|x')$  и апостериорная энтропия источника  $H(x|x')$  будет также равна бесконечности. Количество информации, вносимой этим отсчетом, равно разности энтропий априорного и апостериорного распределений величины  $x$  равно разности соответствующих дифференциальных энтропий, так как второй член в (6.20), одинаковый для обоих случаев, при определении разности энтропий исключается.

Поэтому количество информации, приходящейся в среднем на один отсчет непрерывного сигнала, равно

$$\mathcal{I}(x|x') = H(x) - H(x|x') = h(x) - h(x|x'), \quad (6.22)$$

где

$$h(x|x') = - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w(x, x') \log_2 w(x|x') dx dx'; \quad (6.23)$$

$w(x, x')$  — совместная плотность вероятности значения отчета  $x'$  и фактического значения сигнала  $x$ . Здесь  $h(x)$  и  $h(x|x')$  представляет соответственно априорную (до получения отсчета  $x'$ ) и апостериорную дифференциальные энтропии источника непрерывных сообщений.

Дифференциальная энтропия (6.21) характеризует относительную степень неопределенности различных случайных процессов (распределение источников). В отличие от энтропии дискретного распределения (источника) она может изменяться и даже менять знак в зависимости от масштаба переменной  $x$ . Действительно, изменив масштаб в  $k$  раз, что соответствует переходу к переменной  $x_1 = kx$ , и учитывая, что  $w(x_1) = w(x)/k$ , получим:

$$\begin{aligned}
 h(x_1) &= - \int_{-\infty}^{\infty} w(x_1) \log_2 w(x_1) dx_1 = \\
 &= - \int_{-\infty}^{\infty} \frac{w(x)}{k} \left( \log_2 \frac{w(x)}{k} \right) k dx = - \int_{-\infty}^{\infty} w(x) \log_2 w(x) dx + \\
 &+ \log_2 k \int_{-\infty}^{\infty} w(x) dx = h(x) + \log_2 k.
 \end{aligned}$$

Однако разность значений дифференциальных энтропий двух распределений при одном и том же масштабе переменной  $x$  не зависит от выбора масштаба.

Можно показать [17], что при фиксированной дисперсии случайной величины  $\xi$  дифференциальная энтропия имеет максимальное значение при нормальном распределении этой величины:

$$w(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp \left[ - \frac{(x-a)^2}{2\sigma^2} \right],$$

где  $a$  — математическое ожидание случайной величины  $\xi$ ;  $\sigma^2$  — дисперсия этой величины. Вычислив интеграл (6.21) для этого распределения, получим

$$h_{\text{норм}}(x) = \log_2 \sqrt{2\pi e} \sigma. \quad (6.24)$$

Из формулы (6.24) видно, что дифференциальная энтропия не зависит от математического ожидания  $a$  случайной величины  $x$ .

## Глава 7

ПРОПУСКНАЯ СПОСОБНОСТЬ КАНАЛОВ СВЯЗИ  
И ТЕОРЕМЫ ОПТИМАЛЬНОГО КОДИРОВАНИЯ

## 7.1. Пропускная способность дискретного канала без помех.

## Теорема Шеннона для канала без помех

Среднее количество информации, передаваемой по каналу в единицу времени, называется *скоростью передачи* информации. Если по каналу передается  $v_k$  символов в единицу времени, а среднее количество информации на один символ канала равно  $H(B)$ , то скорость передачи информации по каналу

$$\frac{dI(B)}{dt} = v_k H(B). \quad (7.1)$$

Средняя скорость поступления информации от источника сообщений называется его *производительностью*. Если источник  $A$  сообщений выдает  $v_i$  символов в единицу времени, то его производительность

$$\frac{dI(B)}{dt} = v_i H(A). \quad (7.2)$$

Скорость передачи информации по каналу без помех, определяемая формулой (7.1), зависит как от технических характеристик канала (числа символов, алфавита  $m$  и скорости их передачи  $v_k$ ), так и от статистических свойств входного сообщения (его избыточности). Поэтому для сравнения возможных различных каналов нужно зафиксировать избыточность входного сообщения. Для этой цели удобно использовать безыбыточные входные сообщения, обладающие максимальной энтропией

$$H_{\max}(B) = \log_2 m \quad (7.3)$$

и обеспечивающие максимально возможную для данного канала скорость передачи информации

$$C_k = \left[ \frac{dI(B)}{dt} \right]_{\max} = v_k \log_2 m. \quad (7.4)$$

Величина  $C_k$  характеризующая максимально возможную скорость передачи информации по каналу с данными техническими характеристиками, называется его *пропускной способностью*.

Из формул (7.1) и (7.4) видно, что в общем случае при избыточности входного сообщения, отличной от нуля, скорость передачи информации по каналу меньше его пропускной способности  $C_k$ . Однако, как показал Шеннон, принципиально (оставляя в стороне вопрос технической реализуемости) соответствующим выбором способа кодирования при любой избыточности источника сообщений можно обеспечить скорость передачи информации по каналу без помех, сколь угодно близкую к его пропускной способности, определяемой по формуле (7.4).

Таким образом, пропускная способность канала без помех характеризует не только предельную скорость передачи информации, достижимую при безыбыточном входном сообщении, но и верхнюю границу скорости передачи информации, принципиально достижимую при любых характеристиках источника сообщений. Отсюда следует, что условием согласования источника сообщений с каналом является не соответствие скоростей формирования символов, а соответствие их информационных характеристик: производительности источника и пропускной способности канала.

Для доказательства этого положения введем понятия типичных и нетипичных последовательностей символов и рассмотрим их свойства. Рассмотрим сначала простейший случай, когда алфавит источника сообщений состоит из двух символов  $a_1$  и  $a_2$  вероятности выбора которых соответственно равны  $P_1$  и  $P_2 = 1 - P_1$ , причем вероятность выбора очередного символа не зависит от значений предшествующих символов (символы независимы). Рассмотрим свойства длинных последовательностей таких символов. Вероятность того, что в последовательности из  $n$  символов будет  $r$  символов  $a_1$  и  $n - r$  символов  $a_2$ , определяется биномиальным законом (2.7):

$$P_{r, n-r} = C_n^r P_1^r (1 - P_1)^{n-r} = \frac{n!}{r!(n-r)!} P_1^r (1 - P_1)^{n-r}. \quad (7.5)$$

Здесь  $C_n^r = n!/r!(n-r)!$  — число различных последовательностей, содержащих  $r$  символов  $a_1$  и  $n - r$  символов  $a_2$ . При увеличении общего числа символов  $n$  значения  $r$  и  $n - r$  в каждой реализации последовательности будут стремиться к математическим ожиданиям  $P_1 n$  и  $(1 - P_1)n$  и именно такие последовательности будут типичными. Вероятность появления *нетипичных* последовательностей с другим соотношением символов  $a_1$  и  $a_2$  будет малой. Согласно закону больших чисел [16] при  $n \rightarrow \infty$  вероятность появления типичной последовательности с соотношением символов, строго соответствующим их математическим ожиданиям, стремится к единице,

а вероятность появления нетипичных последовательностей с другим соотношением символов стремится к нулю. При этом все типичные последовательности с  $P_1 n$  символами  $a_1$  и  $(1 - P_1)n$  символами  $a_2$  имеют одинаковую вероятность появления, равную  $P_1^{P_1 n} (1 - P_1)^{(1-P_1)n}$ .

Совершенно аналогичный результат получим для алфавита из  $m$  независимых символов  $a_1, a_2, \dots, a_m$  с вероятностями их появления  $P_1, P_2, \dots, P_m$ . В этом случае типичной последовательностью длительности  $n$  является последовательность, содержащая  $P_1 n$  символов  $a_1, P_2 n$  символов  $a_2$  и т.д.

Число различных типичных последовательностей длительности  $n$  равно

$$\frac{n!}{(P_1 n)! (P_2 n)! \dots (P_m n)!}$$

и все они имеют одинаковую вероятность появления, равную

$$P_{\text{тип}} = P_1^{P_1 n} P_2^{P_2 n} \dots P_m^{P_m n}. \quad (7.6)$$

Полученные результаты можно обобщить и на случай корреляции между символами сообщения, когда также могут быть выделены типичные последовательности с кратностью отдельных символов и их сочетаний, соответствующей статистическим свойствам источника сообщений (во всех случаях, как уже указывалось, предполагается стационарность последовательностей символов, выдаваемых источником). При этом все типичные последовательности по-прежнему равновероятны, а их суммарная вероятность стремится к единице при  $n \rightarrow \infty$ .

Рассмотренные свойства последовательностей символов большой длительности могут быть обобщены следующей теоремой асимптотической равновероятности [17]: *любая реализация последовательностей символов длительностью  $n$ , выдаваемая стационарным источником сообщений, при достаточно большом  $n$  с вероятностью, сколь угодно близкой к единице, совпадает с одной из равновероятных типичных последовательностей.*

Число  $N_{\text{тип}}(A)$  типичных последовательностей достаточно большой длительности  $n$ , создаваемых дискретным источником сообщений  $A$ , может быть выражено через информационную энтропию источника  $H(A)$ . Поскольку при  $n \rightarrow \infty$  источник с вероятностью, близкой к единице, выдает лишь типичные последовательности, имеющие равную вероятность  $1/N_{\text{тип}}(A)$ , то информация, содержащаяся в такой последовательности, согласно (6.16) равна  $\log_2 N_{\text{тип}}(A)$ . С другой стороны суммарная информация,

содержащаяся в  $n$  символах источника с энтропией  $H(A)$  при  $n \rightarrow \infty$  равна  $nH(A)$ . Отсюда получаем, что при  $n \rightarrow \infty$

$$nH(A) = \log_2 N_{\text{тип}}(A)$$

или

$$N_{\text{тип}}(A) = 2^{nH(A)}. \quad (7.7)$$

Согласно (6.19) для источника с объемом алфавита  $m$  и избыточностью  $\rho_n$   $H(A) = (1 - \rho_n) \log_2 m$  и, следовательно,

$$N_{\text{тип}}(A) = 2^{n(1-\rho_n) \log_2 m}. \quad (7.8)$$

Определим долю, которую составляет число типичных последовательностей от общего числа всевозможных последовательностей длиной  $n$  с произвольным сочетанием символов. Общее число всевозможных последовательностей длиной  $n$  при объеме алфавита  $m$  равно

$$N(A) = m^n = 2^{n \log_2 m},$$

откуда

$$\frac{N_{\text{тип}}(A)}{N(A)} = 2^{-n \rho_n \log_2 m}. \quad (7.9)$$

Из (7.9) следует, что при избыточности источника сообщений ( $\rho_n \neq 0$ ) и  $n \rightarrow \infty$ ,  $N_{\text{тип}}(A)/N(A) \rightarrow 0$ . Таким образом, если источник сообщений обладает избыточностью, то с увеличением длительности последовательности символов все меньшая доля от всех возможных последовательностей, составляющая типичные последовательности, участвует в передаче информации. И только в источниках без избыточности при этом число типичных последовательностей совпадает с общим числом возможных последовательностей, и все они в равной мере используются для передачи информации.

Рассмотренные свойства типичных последовательностей позволяют доказать теорему Шеннона для дискретного канала без помех: *если пропускная способность дискретного канала без помех превышает производительность источника сообщений, т.е. удовлетворяется условия*

$$v_n \log_2 m > v_n H(A), \quad (7.10)$$

*то существует способ кодирования и декодирования сообщений источника с энтропией  $H(A)$ , обеспечивающий сколь угодно высокую надежность*

отождествления принятых комбинаций с действительно переданными. Если же  $v_k \log_2 m < v_k H(A)$ , то такого способа нет.

Для доказательства теоремы все типичные последовательности достаточной большой протяженности  $T$  будем кодировать цифровыми кодовыми комбинациями  $B$  той же протяженности  $T$  с основанием  $m$ , равным объему алфавита канала. При этом число символов (разрядность) цифрового кода равно  $Tv_k$  и соответственно число различных кодовых комбинаций  $B$  равно

$$N(B) = m^{Tv_k} = 2^{Tv_k \log_2 m}. \quad (7.11)$$

Число типичных последовательностей длительностью  $T$  с числом символов  $n = Tv_k$  согласно (7.7) равно

$$N_{\text{тип}}(A) = 2^{Tv_k H(A)}. \quad (7.12)$$

Условие (7.10) можно записать в виде  $v_k \log_2 m = v_k H(A) + \epsilon$ , где  $\epsilon$  — сколь угодно малая величина, откуда  $N(B)/N_{\text{тип}}(A) = 2^{\epsilon T}$ .

Если принять  $\epsilon = \log_2 e / TN_{\text{тип}}(A)$ , то

$$\frac{N(B)}{N_{\text{тип}}(A)} = e^{1/N_{\text{тип}}(A)} = 1 + \frac{1}{N_{\text{тип}}(A)} + \frac{1}{2! [N_{\text{тип}}(A)]^2} + \dots$$

или

$$N(B) > N_{\text{тип}}(A) + 1.$$

Таким образом, при выполнении условия (7.10) число различных кодовых комбинаций  $B$  по крайней мере на единицу больше числа типичных последовательностей источника. Эту избыточную кодовую комбинацию поставим в соответствие всем нетипичным последовательностям, предопределив их недостоверную передачу. Поскольку при  $T \rightarrow \infty$  и соответственно  $n \rightarrow \infty$  вероятность появления нетипичной последовательности стремится к нулю, а величина  $\epsilon$ , определяющая требуемое превышение пропускной способности канала над производительностью источника, бесконечно малая, первую часть теоремы можно считать доказанной.

В случае нарушения условия (7.10), когда  $v_k \log_2 m > v_k H(A)$  используя тот же подход, получаем неравенство  $N_{\text{тип}}(A) > N(B) + 1$ .

Таким образом, в этом случае даже при описанном способе кодирования, обеспечивающем равную вероятность использования всех символов алфавита канала и, следовательно, предельную скорость передачи информации по каналу, мы уже не сможем закодировать и передать все

типичные последовательности  $N_{\text{тип}}(A)$ . Поэтому и вторую часть теоремы можно считать доказанной.

Оптимальное кодирование, использованное при доказательстве теоремы Шеннона для дискретного канала без помех, сводится к предельному укрупнению алфавита канала, когда каждый укрупненный символ (кодовая комбинация) отвечает бесконечно длинной последовательности символов источника сообщений. При этом одновременно устраняется корреляция между символами укрупненного алфавита канала и благодаря сохранению только типичных последовательностей обеспечивается равная вероятность их появления. В результате устраняется избыточность сообщения, передаваемого по каналу.

Кодирование способом, изложенным при доказательстве теоремы, связано с задержкой передачи сообщения на время

$$\tau_{\text{зад}} = 2T + T_0, \quad (7.13)$$

где время  $2T$  определяется тем, что кодирование может начаться, когда уже известна вся последовательность символов источника длительностью  $T$ , а декодирование — когда уже принята кодовая комбинация той же длительности;  $T_0$  — время, затрачиваемое на технические операции кодирования, декодирования и на прохождение сигнала по каналу.

Заметим, что в канале без помех источником ненадежности отождествления передаваемых сообщений может быть только операция кодирования, ибо нарушение соответствия при передаче сообщения по каналу исключено.

## 7.2. Скорость передачи информации и пропускная способность дискретного канала при наличии помех

Если в канале без помех среднее количество информации, получаемой с выхода канала в единицу времени, соответствует среднему количеству информации, содержащемуся во входном сообщении той же длительности, то в канале с помехами это соответствие нарушается. При наличии помех скорость передачи информации по каналу будет меньше среднего количества информации, поступающего в единицу времени на вход канала, ибо часть информации, поступившей на вход канала, разрушается помехами. Помехи нарушают взаимно-однозначное соответствие между символами  $b'_j$  на выходе канала и символами  $b_i$  на его входе. Принятый символ  $b'_j$  определяет при этом не конкретный входной символ  $b_i$ , а некоторый закон распределения апостериорных вероятностей  $P(b_i | b'_j)$

возможных значений передававшихся символов. Соответственно апостериорная энтропия входного сообщения после приема конкретного символа  $b'_j$ , которая для канала без помех равна нулю, при наличии помех будет отлична от нуля:

$$H(B|b'_j) = - \sum_{i=1}^m P(b_i|b'_j) \log_2 P(b_i|b'_j). \quad (7.14)$$

Осредняя апостериорную энтропию сообщения по всем возможным значениям принятого символа  $b'_j$  (в общем случае число  $m'$  символов алфавита на выходе канала с помехами может отличаться от числа  $m$  символов входного алфавита), получаем:

$$\begin{aligned} H(B|B') &= - \sum_{j=1}^{m'} \sum_{i=1}^m P(b'_j) P(b_i|b'_j) \log_2 P(b_i|b'_j) = \\ &= - \sum_{j=1}^{m'} \sum_{i=1}^m P(b_i, b'_j) \log_2 P(b_i|b'_j), \end{aligned} \quad (7.15)$$

где  $P(b_i, b'_j)$  — совместная вероятность появления входного символа  $b_i$  и выходного  $b'_j$ .

Используя формулы (2.3) и (2.9), выражение (7.15) можно переписать в виде:

$$\begin{aligned} H(B|B') &= - \sum_{j=1}^{m'} \sum_{i=1}^m P(b_i) P(b'_j|b_i) \times \\ &\times \log_2 \frac{P(b_i) P(b'_j|b_i)}{\sum_{i=1}^m P(b_i) P(b'_j|b_i)}. \end{aligned} \quad (7.16)$$

Величина  $H(B|B')$  по терминологии, введенной К. Шенноном, называется *ненадежностью канала*. Из (7.16) видно, что она зависит как от статистических характеристик входного сообщения канала  $B$  (распределения вероятностей  $P(b_i)$ ), так и от вероятностных характеристик искажений символов, вносимых действующими в канале помехами.

Последние могут быть заданы матрицей условных переходов  $[P_{ij}]$ , где  $P_{ij}$  обозначает условную вероятность  $P(b'_j|b_i)$  трансформации входного символа  $b_i$  в выходной символ  $b'_j$ .

Среднее количество информации, приходящееся на один символ канала с помехами, определяется уменьшением неопределенности знания переданного символа после приема, равным разности априорной и апостериорной энтропий передаваемого сообщения

$$\mathcal{I}(B, B') = H(B) - H(B|B') =$$

$$= H(B) + \sum_{j=1}^{m'} \sum_{i=1}^m P(b_i, b'_j) \log_2 P(b_i|b'_j). \quad (7.17)$$

Наличие двойного аргумента в левой части (7.17) связано с тем, что осреднение информации, приходящейся на один символ, должно производиться по всем сочетаниям символов входного и выходного алфавита.

Преобразуем (7.17), воспользовавшись формулами (2.8) и (2.3):

$$\begin{aligned} \mathcal{I}(B, B') &= - \sum_{i=1}^m P(b_i) \log_2 P(b_i) + \sum_{j=1}^{m'} \sum_{i=1}^m P(b_i, b'_j) \times \\ &\times \log_2 P(b_i|b'_j) = - \sum_{j=1}^{m'} \sum_{i=1}^m P(b_i, b'_j) \log_2 P(b_i) + \\ &+ \sum_{j=1}^{m'} \sum_{i=1}^m P(b_i, b'_j) \log_2 \frac{P(b_i, b'_j)}{P(b'_j)} = \\ &= \sum_{j=1}^{m'} \sum_{i=1}^m P(b_i, b'_j) \log_2 \frac{P(b_i, b'_j)}{P(b_i) P(b'_j)}. \end{aligned} \quad (7.18)$$

В силу симметрии (7.18) относительно символов входного и выходного алфавитов можно заключить, что

$$\mathcal{I}(B, B') = H(B) - H(B|B') = H(B') - H(B'|B) = \mathcal{I}(B', B), \quad (7.19)$$

т.е. среднее количество информации, которое несут один символ выходного сообщения о текущем состоянии входного сообщения и один символ входного сообщения о текущем состоянии выходного сообщения, оказывается одинаковым. Оно может быть выражено как через разность априорной энтропии входного сообщения и апостериорной его энтропии при известных символах выходного сообщения, так и через разность априорной энтропии выходного сообщения и апостериорной его энтропии при известных символах входного сообщения.

Средняя скорость передачи информации по каналу с помехами равна

$$\begin{aligned} \frac{d\mathcal{I}(B, B')}{dt} &= v_k \mathcal{I}(B, B') = v_k [H(B) + \sum_{j=1}^{m'} \sum_{i=1}^m P(b_i, b'_j) \times \\ &\times \log_2 P(b_i|b'_j)]. \end{aligned} \quad (7.20)$$

Под пропускной способностью дискретного канала с помехами, как и в случае канала без помех, будем понимать предельно возможную скорость передачи информации при заданных технических характеристиках канала (числе символов алфавита, скорости их передачи и ненадежности канала). Она соответствует условию максимума среднего количества информации



на один символ на входе канала, т.е. условию безызбыточности входных сообщений:

$$C_k = v_k \mathcal{I}_{\max}(B, B') = v_k [\log_2 m - H(B|B')] . \quad (7.21)$$

В данном случае пропускная способность определяется нами как предельное количество информации, которое достигает выхода канала в единицу времени, если на вход в единицу времени поступает максимальное количество информации  $v_k \log_2 m$ , которое только может быть воспринято каналом с данным алфавитом и скоростью формирования символов.

Пропускная способность дискретного канала с помехами, как следует из (7.21), зависит от ненадежности канала  $H(B|B')$ , характеризующей средние потери информации, приходящейся на один символ, из-за действия помех. При отсутствии помех  $H(B|B') = 0$ , так как

$$\begin{aligned} \text{при } i \neq j \quad P(b_i, b'_j) &= P(b_i|b'_j) = 0, \\ \text{при } i = j \quad P(b_i, b'_j) &= 1. \end{aligned}$$

Пропускная способность канала, определяемая формулой (7.21), в этом случае равна

$$C_{k \max} = v_k \log_2 m,$$

что совпадает с приведенной ранее формулой (7.4) для пропускной способности канала без помех. При полном забитии канала помехами, когда независимо от принятого символа равновероятен любой символ на входе канала, т.е.  $H(B|B') = \log_2 m$ , как следует из (7.21), пропускная способность

$$C_{k \min} = 0.$$

Таким образом, в зависимости от уровня помех (ненадежности канала) пропускная способность канала может менять в пределах

$$0 \leq C_k \leq v_k \log_2 m. \quad (7.22)$$

### 7.3. Основная теорема Шеннона для дискретного канала с помехами

Шеннон доказал, что и в случае канала с помехами его пропускная способность определяет верхнюю границу скорости достоверной передачи информации по каналу.

Теорему Шеннона для этого случая можно сформулировать следующим образом: для дискретного канала с помехами существует такой способ кодирования, при котором может быть обеспечена безошибочная передача всей информации, поступающей от источника сообщений, если только пропускная способность канала превышает производительность источника сообщений, т.е. выполняется условие:

$$v_k [\log_2 m - H(B|B')] > v_n H(A). \quad (7.23)$$

Докажем эту теорему. Число типичных последовательностей достаточно большой длительности  $T$ , которое может давать источник  $A$  производительностью  $v_n H(A)$ , равно

$$N_{\text{тип}}(A) = 2^{T v_n H(A)}. \quad (7.24)$$

Для кодирования этих типичных последовательностей будем применять цифровые коды той же длительности  $T$ , образуемые с использованием символов алфавита канала. При основании цифрового кода  $m$ , равном числу символов алфавита канала, и числе символов в коде  $v_k T$  общее число различных цифровых кодов

$$N(B) = m^{v_k T} = 2^{v_k T \log_2 m}. \quad (7.25)$$

Поскольку  $H(B|B') > 0$ , неравенство (7.23) только усилится, если его записать в виде:

$$v_k \log_2 m > v_n H(A), \quad (7.26)$$

тогда

$$N(B) = 2^{T v_k \log_2 m} \gg 2^{T v_n H(A)} = N_{\text{тип}}(A), \quad (7.27)$$

т.е. множество цифровых кодов  $N(B)$ , которые могут быть использованы для кодирования, много больше числа типичных последовательностей  $N_{\text{тип}}(A)$ , подлежащих кодированию. Таким образом, при кодировании входных типичных последовательностей используется лишь небольшая часть множества  $N(B)$ , что предопределяет большой выбор возможных способов кодирования (число таких способов  $M$ , очевидно, равно числу размещения из  $N(B)$  по  $N_{\text{тип}}(A)$  элементов). Требуется доказать, что среди этих  $M$  способов кодирования имеется по крайней мере один, обеспечивающий однозначную идентификацию любой принятой кодовой комбинации с одной из комбинаций, используемых при кодировании, что является условием достоверного приема. Для доказательства найдем

вероятность правильного приема, осредненную по всем возможным способам кодирования. Поскольку апостериорная энтропия сообщения на входе канала при известных символах на его выходе равна  $H(B|B')$ , то число типичных входных кодовых последовательностей, которые могут трансформироваться в данное выходное сообщение длительностью  $T$ , по аналогии с (7.12) равно

$$N_{\text{тип}}(B|B') = 2^{T v_k H(B|B')} . \quad (7.28)$$

Правильный прием (однозначная идентификация переданной кодовой комбинации) обеспечивается в том случае, когда среди  $N_{\text{тип}}(B|B')$  входных последовательностей, которые могли дать данную выходную последовательность, лишь одна была использована при кодировании, а остальные  $N_{\text{тип}}(B|B') - 1$  последовательностей вообще не могли передаваться.

Найдем осредненную по всем возможным способам кодирования вероятность  $\bar{P}_{\text{прав}}$  того, что ни одна из этих  $[N_{\text{тип}}(B|B') - 1]$  последовательностей не была использована при кодировании. Средняя [для всех возможных способов кодирования множества  $N_{\text{тип}}(A)$  поступающих от источника типичных последовательностей] вероятность использования конкретной кодовой комбинации из  $N(B)$  возможных комбинаций равна

$$P_{\text{тип}} = N_{\text{тип}}(A)/N(B) ,$$

ибо осреднению по всем возможным способам кодирования соответствует равновероятный выбор кодовых комбинаций для кодирования каждой входной последовательности. Соответственно средняя вероятность того, что конкретная кодовая комбинация не была использована при кодировании, равна

$$1 - P_{\text{тип}} = 1 - N_{\text{тип}}(A)/N(B) . \quad (7.29)$$

Средняя же вероятность того, что не использовались конкретные  $[N_{\text{тип}}(B|B') - 1]$  кодовые комбинации, определяющие среднюю вероятность правильного приема, равна

$$\bar{P}_{\text{прав}} = (1 - P_{\text{тип}})^{[N_{\text{тип}}(B|B') - 1]} = \left[ 1 - \frac{N_{\text{тип}}(A)}{N(B)} \right]^{[N_{\text{тип}}(B|B') - 1]} . \quad (7.30)$$

Так как основание степени в правой части (7.30) меньше единицы, то, увеличивая показатель степени до значения  $N_{\text{тип}}(B|B')$ , из (7.30) получаем неравенство

$$\bar{P}_{\text{прав}} > \left[ 1 - \frac{N_{\text{тип}}(A)}{N(B)} \right]^{N_{\text{тип}}(B|B')} ,$$

а после разложения в ряд правой части этого неравенства имеем:

$$\begin{aligned} \bar{P}_{\text{прав}} &> 1 - N_{\text{тип}}(B|B') \frac{N_{\text{тип}}(A)}{N(B)} + \\ &+ \frac{N_{\text{тип}}(B|B') [N_{\text{тип}}(B|B') - 1]}{2} \left[ \frac{N_{\text{тип}}(A)}{N(B)} \right]^2 - \dots \end{aligned} \quad (7.31)$$

В силу (7.27)  $N_{\text{тип}}(A)/N(B) \ll 1$  и правая часть (7.31), представляющая знакочередующийся ряд с убывающими по абсолютной величине членами, только уменьшится, если отбросить все остальные члены, начиная с положительного члена 2-го порядка. Поэтому можем записать

$$\bar{P}_{\text{прав}} > 1 - \frac{N_{\text{тип}}(B|B') N_{\text{тип}}(A)}{N(B)} . \quad (7.32)$$

Используя (7.24), (7.25), (7.28), из (7.32) получаем:

$$\begin{aligned} \bar{P}_{\text{ош}} &= 1 - \bar{P}_{\text{прав}} < N_{\text{тип}}(B|B') \frac{N_{\text{тип}}(A)}{N(B)} = \\ &= 2^{-T[v_k \log_2 m - v_k H(B|B') - v_k H(A)]} = 2^{-T[C_k - v_k H(A)]} . \end{aligned} \quad (7.33)$$

Помимо ошибки при приеме типичных последовательностей, вероятность которой  $\bar{P}_{\text{ош}}$  определяется формулой (7.33), может также иметь место нарушение достоверности передачи информации при кодировании всякий раз, когда потребуется передача нетипичных последовательностей, кодирование которых вообще не предусматривается. При вероятностях появления нетипичной последовательности  $\delta$  и типичной последовательности  $1 - \delta$  суммарная вероятность ошибочной передачи информации равна

$$\bar{P}_{\Sigma} = \delta + (1 - \delta) \bar{P}_{\text{ош}} . \quad (7.34)$$

При  $T \rightarrow \infty$  в соответствии с (7.33)  $\bar{P}_{\text{ош}} \rightarrow 0$  [по условию теоремы  $C_k > v_k H(A)$ ] и в соответствии с теоремой асимптотической равновероятности  $\delta \rightarrow 0$  следовательно, и  $\bar{P}_{\Sigma} \rightarrow 0$ . Итак, при любом заданном  $\varepsilon > 0$  можно выбрать такое  $T$ , что будет обеспечиваться  $\bar{P}_{\Sigma} < \varepsilon$ .

Поскольку среди всех  $M$  вариантов кодирования обязательно существует хотя бы один, для которого вероятность ошибки не превышает среднего по всем вариантам кодирования значения  $\bar{P}_{\Sigma}$ , теорему Шеннона можно считать доказанной. Заметим, что здесь, как и в канале без помех [см. (7.31)], кодирование связано с задержкой передачи сообщения на время  $\tau_{\text{зад}} = 2T + T_0$ .

Теорема Шеннона для канала с шумами не указывает конкретного способа кодирования, обеспечивающего достоверную передачу информации

со скоростью, сколь угодно близкой к пропускной способности канала, а лишь доказывает принципиальное существование такого кода. Однако из формулы (7.33) вытекает крайне важный практический вывод: достоверность связи тем выше, чем больше длительность кодированной последовательности (при этом соответственно увеличивается задержка в приеме информации) и чем менее эффективно используется пропускная способность канала [больше "запас" по пропускной способности  $C_k - v_k H(A)$ ]. Таким образом, существует возможность "размена" эффективности использования канала на достоверность передачи информации, что и делается на практике.

Теорема Шеннона для канала с помехами имела огромное значение для становления правильных воззрений на принципиальные возможности техники связи. До К. Шеннона считалось, что сколь угодно малую вероятность ошибки можно получить лишь при стремлении скорости передачи информации к нулю. Введя в рассмотрение кодирование последовательностей бесконечной длительности, К. Шеннон впервые показал, что принципиально существуют коды, которые обеспечивают сколь угодно малую вероятность ошибки при конечной скорости передачи информации, причем эти коды обладают сравнительно небольшой избыточностью. Действительно, среднее количество информации, приходящееся на один символ канала при выполнении условия (7.23), равно

$$H(B) = \frac{v_k H(A)}{v_k} = \frac{v_k [\log_2 m - H(B|B')] - \varepsilon]}{v_k} \cong \log_2 m - H(B|B')$$

и, следовательно, избыточность входного сообщения

$$\rho_k = 1 - \frac{H(B)}{\log_2 m} \cong \frac{H(B|B')}{\log_2 m}. \quad (7.35)$$

Таким образом, для обеспечения достоверной передачи информации достаточно ввести во входное сообщение избыточность, превышающую потерю информации в канале из-за действия помех, характеризующую ненадежностью канала  $H(B|B')$ . Удлинение же кодируемых последовательностей при использовании кодов Шеннона вызывает не снижение пропускной способности, а лишь увеличение задержки в приеме информации.

Основной смысл теоремы Шеннона сводится к следующему. Пропускную способность дискретного канала мы определяли как максимальную скорость передачи информации по каналу, соответствующую безызбыточности входных сообщений, т.е. условию поступления на вход канала в каждую единицу времени предельного количества информации, которое только способен воспринять канал. При отсутствии помех  $C_{k \max} = v_k \log_2 m$ , а при наличии помех — меньше на величину  $v_k H(B|B')$ , характеризующую

скорость разрушения информации помехами в канале. Здесь пропускная способность представляет собой чисто техническую характеристику канала как переносчика символов. Вопрос достоверности передачи всей информации, поступающей от источника, при этом остается в стороне. Под пропускной способностью канала можно понимать также предельную скорость *достоверной* передачи по каналу информации от произвольного источника сообщений, обеспечиваемую при оптимальном для данного источника способе кодирования. Это уже характеристика канала в широком смысле, включая кодирующие и декодирующие устройства, увязанная с его назначением. К. Шеннон доказал, что теоретически оба эти определения тождественны: пропускная способность, определяемая как чисто техническая характеристика канала для некоторого типового (безызбыточного) входного сообщения, является и верхней границей скорости достоверной передачи информации от произвольного источника сообщений при оптимальном кодировании бесконечно длинных последовательностей символов. И хотя для реальных систем эта граница недостижима, так как кодирование бесконечно длинных последовательностей символов технически нереализуемо, она служит той идеальной мерой, по которой можно судить о степени совершенства реальных систем. По мере увеличения возможностей дискретной техники разрыв между скоростью передачи информации при требуемой ее достоверности, обеспечиваемой реальными системами, и пропускной способностью канала будет уменьшаться, но ликвидировать его можно было бы только, используя идеальные коды Шеннона, что технически всегда останется неосуществимым.

Для наглядного пояснения роли теорем Шеннона прибегнем к следующему сравнению. Пусть имеется трубопровод для доставки некоторого жидкого продукта, поступающего от источника. Технические возможности трубопровода определяются количеством жидкости, которое можно передать по нему в единицу времени. Скорость передачи продукта по трубопроводу будем определять количеством чистого продукта, доставляемого потребителю в единицу времени, а производительность источника — количеством чистого продукта, поступающего от него в единицу времени. Пропускную способность трубопровода определим как максимально возможную скорость передачи чистого продукта, соответствующую условию, что от источника поступает чистый продукт без примесей (без "избыточности"). Аналогом радиоканала с помехами может служить трубопровод с утечкой. Пропускная его способность, отвечающая скорости доставки продукта потребителю при поступлении на вход трубопровода чистого продукта, будет меньше, чем в трубопроводе без утечки, на величину утечки продукта за единицу времени. Можно теперь представить, какой эффект вызвало бы утверждение, что существует такой способ введения примеси ("избыточности") в продукт, при котором, введя количество примеси, равное утечке в трубопроводе, можно по нему доставить чистый продукт без потерь со скоростью, отвечающей пропускной способности трубопровода с утечкой. Именно такой смысл применительно к задаче передачи информации имеет теорема Шеннона для радиоканала с помехами.

### 7.4. Пропускная способность непрерывного канала при наличии аддитивного шума

Под пропускной способностью непрерывного канала будем понимать максимально возможную скорость передачи информации по каналу при заданных технических характеристиках канала: полосе пропускания и отношении сигнал/шум. Так же, как и в случае дискретных каналов, скорость передачи информации по непрерывному каналу зависит от выбора ансамбля входных сигналов. Максимальная скорость передачи информации реализуется при максимальной скорости поступления информации на вход канала, которая соответствует случаю выбора ансамбля входных сигналов с максимальной энтропией. Для дискретных каналов, когда заданным для ансамбля сигналов является объем алфавита канала, максимальная скорость поступления информации обеспечивается при равномерном использовании символов алфавита. Для непрерывных сигналов, когда заданной следует считать среднюю мощность сигнала, максимальная скорость поступления информации соответствует использованию нормальных центрированных случайных сигналов. Условие центрированности обеспечивает при этом максимум дисперсии при данной средней мощности сигнала, а его нормальность — наибольшую априорную энтропию каждого отсчета при данном значении дисперсии (см. § 6.3). Поэтому при определении пропускной способности непрерывного канала входные сигналы будем считать стационарными случайными центрированными нормальными.

Шум будем считать нормальным белым, что соответствует максимальной энтропии, вносимой шумом в полученное непрерывное сообщение, т.е. максимальной ненадежности канала по Шеннону при заданном отношении сигнал/шум. Принимаемый сигнал имеет вид

$$\eta(t) = \xi(t) + n(t), \quad (7.36)$$

где  $\xi(t)$  — передаваемый сигнал, несущий полезную информацию;  $n(t)$  — шум в канале.

Ширина спектра сигнала ограничивается верхней частотой  $\mathcal{F}_k$ , определяемой полосой пропускания канала. Поэтому согласно теореме Котельникова сигнал полностью определится дискретными отсчетами  $\eta_i$ , следующими с частотой  $2\mathcal{F}_k$ . Значения белого шума при такой частоте отсчетов будут некоррелированными (см. § 2.8). Дискретные отсчеты передаваемого сигнала будем также считать некоррелированными, что соответствует условию максимума его энтропии (для нормальных сигналов, как было показано в § 2.6, некоррелированность отсчетов означает и их

независимость). Тогда и дискретные отсчеты принимаемого сигнала будут некоррелированными (а значит, и независимыми), и количество информации, содержащейся в принимаемом сигнале, будет равно сумме количеств информации, содержащихся в независимых его отсчетах, следующих с частотой  $2\mathcal{F}_k$ . Количество информации о текущем значении передаваемого сигнала  $\xi_i$ , вносимое дискретным отсчетом принимаемого сигнала  $\eta_i$ , может быть представлено в виде разности априорной энтропии этого отсчета и апостериорной его энтропии при известном отсчете передаваемого сигнала (см. § 7.2). Заменяя разность энтропий разностью соответствующих дифференциальных энтропий по аналогии с (7.19), получаем

$$\mathcal{I}(\eta_i, \xi_i) = h(\eta_i) - h(\eta_i/\xi_i). \quad (7.37)$$

Отсчеты принимаемого сигнала, представляющего сумму двух нормальных случайных сигналов  $\xi(t)$  и  $n(t)$ , распределены по нормальному закону с дисперсией  $\sigma_\eta^2 = \sigma_\xi^2 + \sigma_n^2$  (входной сигнал и шум считаем независимыми). Апостериорная неопределенность принимаемого сигнала при известных значениях отсчетов передаваемого сигнала определяется значением шума, имеющего нормальное распределение с дисперсией  $\sigma_n^2$ . Поэтому разность дифференциальных энтропий в правой части (7.37) в соответствии с (6.24) равняется  $\log_2(\sigma_\eta/\sigma_n)$  и информация, вносимая одним дискретным отсчетом сигнала,

$$\begin{aligned} \mathcal{I}(\eta_i, \xi_i) &= \log_2(\sigma_\eta/\sigma_n) = \\ &= \log_2 \sqrt{(\sigma_\xi^2 + \sigma_n^2)/\sigma_n^2} = \log_2 \sqrt{1 + \sigma_\xi^2/\sigma_n^2}. \end{aligned} \quad (7.38)$$

Величина  $\sigma_\xi^2/\sigma_n^2 = \mathcal{P}_s/\mathcal{P}_m$  характеризует отношение сигнал/шум в канале.

Пропускная способность канала, определяемая количеством информации, передаваемой по каналу в единицу времени при рассмотренных условиях, равна информации, содержащейся в  $2\mathcal{F}_k$  независимых дискретных отсчетах сигнала:

$$C_k = 2\mathcal{F}_k \log_2 \sqrt{1 + \mathcal{P}_s/\mathcal{P}_m} = \mathcal{F}_k \log_2(1 + \mathcal{P}_s/\mathcal{P}_m). \quad (7.39)$$

Формула (7.39) называется формулой Шеннона.

Если распределение аддитивной помехи не является нормальным или же спектр ее неравномерный (шум не белый, и имеется корреляция между его отсчетами), то формула (7.39) занижает пропускную способность канала.

Нетрудно видеть, что при  $\mathcal{P}/\mathcal{P}_\text{ш} \rightarrow 0$ ,  $C_\kappa \rightarrow 0$ . Формулу (7.39) можно трактовать следующим образом. Величина  $(1 + \mathcal{P}/\mathcal{P}_\text{ш})$  характеризует количество уровней непрерывного сигнала, различимых на фоне шума при данном отношении  $\mathcal{P}/\mathcal{P}_\text{ш}$ . Поэтому количество информации, приходящееся на один отсчет, будет в данном случае таким же, как для дискретного источника с числом состояний  $(1 + \mathcal{P}/\mathcal{P}_\text{ш})$ .

Преобразуем (7.39), приняв  $\mathcal{P}_\text{ш} = N_0 \mathcal{F}_\kappa$ , где  $N_0$  — спектральная плотность белого шума (для одностороннего спектра):

$$C_\kappa = \mathcal{F}_\kappa \log_2 (1 + \mathcal{P}/N_0 \mathcal{F}_\kappa). \quad (7.40)$$

Раскрывая по правилу Лопиталья неопределенность при  $\mathcal{F}_\kappa \rightarrow \infty$ , в правой части (7.40) получаем  $C_\kappa = (\mathcal{P}/N_0) \log_2 e$ , т.е. пропускная способность стремится к фиксированной величине, определяемой отношением средней мощности сигнала к спектральной плотности шума. Таким образом, обмен мощности сигнала на полосу пропускания для обеспечения заданной пропускной способности непрерывного канала возможен лишь в некоторых границах, за которыми дальнейшее расширение полосы пропускания дает уже малый эффект.

Можно показать, что и в случае непрерывного канала его пропускная способность  $C_\kappa$  является одновременно и верхней границей скорости достоверной передачи информации. Примем следующую модель кодирования и декодирования непрерывных сигналов. Из всего множества реализаций случайного сигнала  $\xi(t)$  будем использовать для передачи счетное подмножество реализаций  $\{x_i(t)\}$ . Кодирование заключается в отождествлении данной реализации  $x(t)$  случайного сигнала  $\xi(t)$  с ближайшим к ней сигналом подмножества  $\{x_i(t)\}$ , а декодирование — в выделении сигнала подмножества  $\{x_i(t)\}$ , ближайшего к принятой реализации выходного сигнала  $\eta(t)$ . Оптимальное кодирование должно обеспечивать достоверный прием переданных сигналов  $x_i(t)$  при максимально возможном числе реализаций  $x_i(t)$ , используемых при кодировании. Правильный прием переданного сигнала  $x_i(t)$  будет обеспечиваться в случае, если из всего подмножества  $\{x_i(t)\}$  ближайшим к принятой реализации случайного сигнала  $\eta(t) = x_i(t) + n(t)$  останется сигнал  $x_i(t)$ . Это накладывает ограничения на минимально допустимое расстояние между реализациями  $x_i(t)$  [при данном уровне шума  $n(t)$ ], следовательно, на максимальное число  $N_{\text{max}}$  сигналов, используемых для передачи информации, и на максимальную скорость достоверной передачи информации.

$$C_{\text{дост}} = (\log_2 N_{\text{max}}) / T,$$

где  $T$  — длительность сигналов.

В [18, 23] доказывается, что скорость достоверной передачи информации по непрерывному каналу при рассмотренном способе кодирования ограничивается пропускной способностью непрерывного канала, определяемой формулой (7.39). При этом так же, как в дискретных каналах, достоверная передача информации со скоростью, приближающейся к пропускной способности непрерывного канала, возможна лишь при кодировании сигналов очень большой (строго говоря, бесконечной) длительности.

## III ОСНОВЫ ТЕОРИИ КОДИРОВАНИЯ

### Глава 8 ОБЩИЕ СВЕДЕНИЯ О КОДИРОВАНИИ

#### 8.1. Функциональная схема радиолинии передачи дискретных сообщений

Общая функциональная схема радиолинии передачи сообщений включает в себя кодирующее и передающее устройства, линию связи, приемное и декодирующее устройства (рис. 8.1). Сообщение, подлежащее передаче, поступает от источника сообщений на вход кодирующего устройства. Операция кодирования заключается в отображении по определенным правилам поступающего от источника сообщения в последовательность символов канала. Правила кодирования должны обеспечивать требуемую достоверность передачи информации за счет введения соответствующей избыточности в сообщение, передаваемое по каналу.

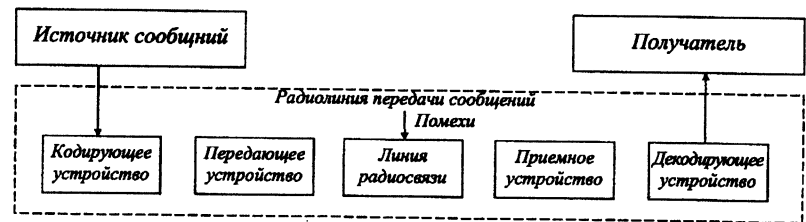


Рис. 8.1

Символы канала представляют электрические сигналы с определенной комбинацией некоторых признаков, называемых кодовыми. Кодовыми признаками могут служить уровень или полярность постоянного напряжения или тока, амплитуда, фаза или частота синусоидального сигнала, амплитуда, длительность или временная расстановка импульсов и т.д. Таким образом, кодирование есть процесс преобразования сообщений в последовательность электрических сигналов, имеющих определенную совокупность тех или иных кодовых признаков. Такие сигналы будем называть *кодовой комбинацией*. Правило, согласно которому каждому сообщению однозначно ставится в соответствие определенная кодовая комбинация,

называется *кодом*. В некоторых случаях, когда источник выдает сообщения в форме электрических сигналов и нет необходимости в изменении их избыточности для передачи по каналу, потребность в кодировании как в специальной операции может отпасть. Однако в большинстве случаев, особенно в радиоприемах с повышенными требованиями к достоверности передачи информации, операция кодирования необходима.

Передатчик служит для преобразования электрического сигнала, представляющего кодовую комбинацию, в колебание высокой частоты, используемое для передачи по линии радиосвязи. В передатчике осуществляется модуляция (или манипуляция) параметров высокочастотного колебания (амплитуды, фазы или частоты) по закону, отображающему признаки передаваемой кодовой комбинации. От вида модуляции зависит различимость признаков кодовой комбинации на фоне помех и, следовательно, помехоустойчивость канала. Под линией радиосвязи (в узком смысле) понимают среду, в которой распространяется вырабатываемый передатчиком высокочастотный сигнал. В процессе передачи сигнала по линии радиосвязи на него могут действовать помехи, под которыми понимаются все сторонние возмущения, препятствующие правильному приему сигнала. Помехи могут действовать и внутри аппарата (шумы приемника). В случае действия аддитивных помех их удобно приводить ко входу приемника.

Приемник обеспечивает прием высокочастотного сигнала и его демодуляцию. На выходе приемника получается последовательность кодовых комбинаций, которая в общем случае (из-за действия помех) может отличаться от переданных комбинаций. Способность приемника минимизировать возникающие при действии помех искажения характеризуют его *помехоустойчивость*.

Декодирующее устройство осуществляет обратное преобразование кодовых комбинаций в сообщения. Из-за действия помех возможно нарушение однозначного соответствия между переданным и принятым сообщением (ошибочный прием). Достоверность передачи информации зависит от формы сигнала и характеристик радиоприема. Форма сигнала определяется кодом и видом модуляции. От вида модуляции зависит вероятность искажения элементарных символов (признаков) кодовой комбинации. От выбора кода зависит вероятность правильного приема сообщения при данной вероятности искажения элементарных символов. Таким образом, помехозащищенность радиоприема передачи сообщений зависит от выбора как способа кодирования (кода), так и вида модуляции.

С выхода декодирующего устройства принятое сообщение поступает к получателю.

## 8.2. Цифровые коды

Передачу дискретных сообщений всегда можно свести к передаче соответствующий чисел. В случае, когда информация, содержащаяся в сообщении, дает количественную характеристику, она, естественно, выражается числом. Если же информация отражает качественно разные состояния (явления и т.п.), то пронумеровав их, можно свести передачу сообщений к передаче чисел, определяющих порядковый номер каждого состояния. В частности, при передаче текста, можно пронумеровать буквы алфавита и передавать числа (код Бодо).

Число различных элементарных символов определяется основанием цифрового кода; оно может быть на единицу меньше основания кода, если один из символов обозначается отсутствием сигнала (код с пассивным нулем).

Увеличение основания цифрового кода приводит к повышению пропускной способности радиоприема (уменьшается требуемое число разрядов кода), но зато требует большего числа различных элементарных символов, что приводит к усложнению радиоприема. При отсутствии помех пропускная способность канала увеличивалась бы пропорционально увеличению основания кода. Наличие помех ослабляет этот эффект: увеличение числа различных символов, если энергия их остается фиксированной, т.е. сохраняется скорость их формирования, ухудшает их различимость на фоне помех и повышает вероятность их искажений. Это приводит к увеличению ненадежности канала и снижает его пропускную способность (см. § 7.2). Эффект повышения пропускной способности канала в результате увеличения основания кода будет тем меньшим, чем большим будет уровень помех.

Наибольшее распространение среди цифровых кодов получил двоичный код, отличающийся простотой технической реализации.

## 8.3. Основные задачи теории кодирования

Шенноновскую теорию идеального кодирования, основывающуюся на кодировании бесконечно длинных последовательностей символов, трудно непосредственно применить в практике связи по следующим причинам:

1. Реальные сообщения имеют конечную длительность.
2. Число различных кодовых комбинаций с увеличением их длительности, как следует из (7.11) и (7.25), растет по экспоненте. Поэтому приближение к условиям идеального кодирования по Шеннону в результате

увеличения длительности кодов ведет к практически неприемлемому усложнению кодирующих и декодирующих устройств, а также устройств памяти.

3. Кодирование и декодирование по Шеннону может осуществляться только после накопления в памяти всей последовательности символов длительностью  $T$ , что требует дополнительной задержки в передаче информации на время  $2T$ . Это также ограничивает приемлемую длительность кодов и возможность приближения к условиям идеального кодирования.

Теоремы Шеннона играют роль предельных теорем. Они послужили толчком к бурному развитию прикладной теории кодирования, имеющей своей целью оптимизацию технически реализуемых кодов ограниченной длительности. Решение этих задач требует качественно иного подхода, ибо сообщения ограниченной длительности уже не обладают присущими бесконечно длинным последовательностям символов асимптотическими свойствами, на которых основывается вся теория идеального кодирования Шеннона. При этом, если для каналов без помех еще имеется принципиальная возможность при ограниченной длительности кодовых комбинаций приблизиться к характеристикам канала, обеспечиваемым при оптимальном кодировании по Шеннону, то для каналов с помехами технически реализуемые способы кодирования принципиально не могут обеспечить достоверную передачу информации со скоростью, определяемой пропускной способностью канала по Шеннону.

Исследованием свойств кодов ограниченной длительности и поиском оптимальных структур технически реализуемых кодов занимается теория кодирования. Теория кодирования развивается в двух главных направлениях: 1) поиски кодов, позволяющих в каналах без помех максимально удержать избыточность источника и тем самым повысить эффективность передачи информации (теория экономного кодирования); 2) поиски кодов, повышающих достоверность передачи информации в каналах с помехами (теория помехоустойчивого кодирования).

Рассмотрим кратко пути решения этих основных задач теории кодирования.

## 8.4. Понятие об экономном кодировании

Хотя оптимальное кодирование, использовавшееся при доказательстве теоремы Шеннона для дискретных каналов без помех, технически не реализуемо, близкие к шенноновскому пределу результаты можно получить и для последовательностей символов ограниченной длительности,

используя так называемое статистическое кодирование. Как известно, максимальная скорость передачи информации, отвечающая пропускной способности канала без помех, реализуется при формировании на входе канала безыбыточных сообщений, отвечающих условию независимости и равной вероятности использования всех символов алфавита канала. Физическую сущность достижения при этих условиях максимальной скорости передачи информации можно трактовать как постоянство скорости передачи информации на элементарных интервалах, отводимых на передачу каждого символа (все символы несут одинаковую информацию и имеют одинаковую длительность). Однако этого можно добиться и при использовании алфавита с неравновероятными, но независимыми символами, если время отводимое на передачу отдельных символов, принимать пропорциональным доставляемой ими информации, т.е. выбрать из условия

$$t_j = -(\log_2 P_j)/C_k, \quad (8.1)$$

где  $C_k$  — пропускная способность канала.

Кодирование, при котором обеспечивается распределение времени на передачу отдельных символов алфавита в зависимости от априорных вероятностей их появления по закону (8.1), называется *оптимальным статистическим кодированием*.

Среднее время передачи одного символа, при котором обеспечивается пропускная способность канала  $C_k$ , в случае оптимального статистического кодирования равно

$$t_{cp} = \sum_{j=1}^m P_j t_j = -\frac{1}{C_k} \sum_{j=1}^m P_j \log_2 P_j = \frac{H(A)}{C_k}, \quad (8.2)$$

т.е. соответствует времени на передачу одного символа, отвечающему той же пропускной способности при идеальном кодировании по Шеннону.

Необходимой предпосылкой применения оптимального статистического кодирования является независимость символов алфавита, обеспечивающая априорное (независимое от контекста) знание вероятностей  $P_j$  появления отдельных символов. Реальные сообщения, подлежащие передаче, обычно не удовлетворяют условию независимости символов. Поэтому статистическому кодированию должны предшествовать меры по устранению или, по крайней мере, по ослаблению корреляционных связей между символами алфавита канала. Этого можно достигнуть, используя укрупнение алфавита канала, которое сводится к тому, что блок из  $L$  символов источника сообщений объединяется в один символ алфавита канала. При этом практически устраняются статистические связи между символами укрупненного алфавита канала. При объеме алфавита источника, равном  $m$ , объем укрупненного алфавита  $m_L$  при любых возможных сочетаниях



символов источника в блоке длины  $L$  равен  $m_y = m^L$ . Уменьшения избыточности при этом практически не происходит, поскольку при кодировании относительно коротких последовательностей еще не происходит выделения существенной группы нетипичных последовательностей, вероятностью появления которых можно было бы пренебречь, исключив их из укрупненного алфавита. А без этого, как легко убедиться, изменить избыточность алфавита нельзя. При энтропии источника сообщений  $H(A)$  среднее количество информации на символ укрупненного фильтра равно  $H_y = LH(A)$  и избыточность передаваемого сообщения

$$\rho_{\text{из}} = 1 - \frac{H_y}{H_{\text{ymax}}} = 1 - \frac{LH(A)}{\log_2 m_y} = 1 - \frac{LH(A)}{L \log_2 m} = 1 - \frac{H(A)}{\log_2 m} = \rho_{\text{и}},$$

т.е. равна избыточности сообщения, поступающего от источника. Эта избыточность вызвана большим различием вероятностей появления символов укрупненного алфавита (в отличие от рассматриваемых Шенноном очень длинных последовательностей, когда можно считаться только с равновероятными типичными последовательностями, составляющими лишь небольшую часть из всех возможных сочетаний символов данной длительности).

Распределение длительностей символов, близкое к закону (8.1), можно реализовать, используя для их передачи неравномерные цифровые коды, т.е. цифровые коды, содержащие различное число разрядов. При этом строгое удовлетворение условия (8.1) не обеспечивается из-за дискретного закона изменения длительности цифрового кода. Однако, чем больше длина  $L$  кодируемых блоков и соответственно объем укрупненного алфавита  $m_y$ , тем большей будет максимальная длительность цифрового кода и меньшим влияние дискретности ее изменения.

Таким образом, используя для каналов без помех статистические методы кодирования, можно получить результат, близкий к идеальному кодированию по Шеннону, ибо при практически приемлемой длительности  $L$  кодируемых последовательностей можно достигнуть весьма малых отклонений от оптимальности, обусловленных как неполным разрушением корреляционных связей между символами укрупненного алфавита, так и дискретностью изменения длительности используемых для кодирования цифровых кодов.

Рассмотрим в качестве конкретного примера статистического кодирования двоичный код Шеннона—Фано. Для его формирования расположим все  $m_y$  независимых символов укрупненного алфавита канала в порядке возрастания априорных вероятностей их появления  $P_j$ . На рис. 8.2 показан случай использования восьми символов, обозначенных первыми восемью буквами латинского алфавита; в скобках указаны вероятности появления этих символов. Найдем сечение, разбивающее все символы на две группы

с суммарной вероятностью каждой примерно равной 0,5. На рис. 8.3 это сечение проходит между символами  $f$  и  $g$ . Присвоим для всех символов, входящих в первую группу, единицу в передаваемом первом разряде отображающих их двоичных кодов, а для символов второй группы — соответственно нуль. Разобьем снова каждую из этих групп на две подгруппы с суммарными вероятностями примерно равными 0,25 и присвоим следующим разрядам двоичных кодов, отображающих символы первых подгрупп каждой группы, единицы, а вторых подгрупп — нули. Кодирование

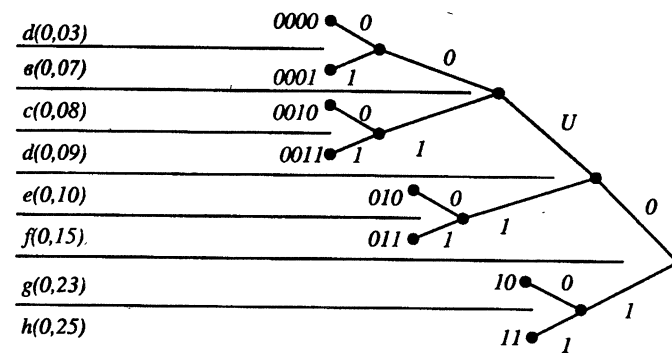


Рис. 8.2

каждого символа заканчивается, как только он остается единственным символом в подгруппе после соответствующего этапа разбиений. Очевидно, что чем меньше вероятность появления символа, тем в большем числе разбиений он будет участвовать и тем длиннее будет отображающий его двоичный код. Процедуру кодирования удобно осуществлять, используя построение "кодowego дерева", как показано на рис. 8.3. Полученные двоичные коды, отображающие каждый из символов, записаны на рисунке у соответствующих вершин кодowego дерева. Это неравномерные двоичные коды, обладающие тем свойством, что ни одна из используемых кодовых комбинаций меньшей длины не совпадает с началом кодовой комбинации большей длины. Такие неравномерные коды называются *неприводимыми*. В отсутствие искажений символов в канале неприводимость кодов обеспечивает однозначность декодирования неравномерных кодов без использования специальных символов для разделения кодовых комбинаций: конец каждой кодовой комбинации и начало следующей в этом случае определяются единственным образом. Заметим, что для обеспечения этого свойства кодовые комбинации обязательно должны записываться, начиная от корня кодowego дерева.



Отклонение распределения длительностей полученных кодовых комбинаций от оптимального закона (8.1) будет тем меньше, чем более равными оказывались суммарные вероятности подгрупп символов при каждом разбиении. В случае равенства суммарных вероятностей подгрупп при всех разбиениях код Шеннона—Фано обеспечивает оптимальное статистическое кодирование.

Экономное кодирование можно принять лишь в случае, когда можно не считаться с влиянием помех (либо из-за малого их уровня, либо в силу низких требований к достоверности передачи информации). При наличии же помех экономное кодирование приводит к значительно большей потере информации на один ошибочно принятый элементарный символ кода, чем кодирование без укрупнения алфавита (теряется не один символ источника сообщений, а целый блок символов длиной  $L$ ).

## 8.5. Общая характеристика задачи помехоустойчивого кодирования

Рассмотрим общую постановку задачи помехоустойчивого кодирования при ограниченной длительности кодируемых сообщений. От источника сообщений поступают последовательности символов  $b_1, b_2, \dots$ , которые называются информационными последовательностями. При кодировании каждой информационной последовательности  $b_1, b_2, \dots$  приводится во взаимно-однозначное соответствие передаваемое сообщение (кодовая последовательность)  $\beta_1, \beta_2, \dots$ . В результате действия помех на выходе канала будет принято сообщение  $\beta'_1, \beta'_2, \dots$  часть символов которого может отличаться от символов передаваемого сообщения:  $\beta'_v \neq \beta_v$ . Если при приеме каждый символ обязательно идентифицируется с одним из символов алфавита канала  $B$  (канал без стирания), т.е. всегда  $\beta'_v \in B$ , то единственной формой ошибки может быть трансформация символов. Если же по некоторым из принятых символов решение об их идентификации с определенным символом алфавита  $B$  не принимается и они оказываются вовсе не опознанными (канал со стиранием), то могут иметь место два вида ошибок: трансформация символов и неопознание (стирание) символа. Неопознанные (стираемые) символы обозначим  $\Phi$ , где  $\Phi \notin B$ .

Помехоустойчивое кодирование сообщений имеет задачей обеспечить с достаточно высокой вероятностью либо обнаружение факта наличия ошибок в принятых сообщениях, либо обнаружение и исправление этих ошибок. Соответственно различают коды с обнаружением ошибок и коды с исправлением ошибок. Код с обнаружением ошибок обеспечивает стирание или особую отметку той части сообщения, в которой обнаружены

ошибки. Такой код эффективен в системах с обратными информационными связями, когда имеется возможность повторно запросить часть сообщения, в котором обнаружены ошибки. При отсутствии обратных информационных связей код с обнаружением ошибок эффективен лишь в таких системах, которые менее критичны к пропуску (стиранию) части сообщения, чем к неверному его приему. Код с исправлением ошибок позволяет получить верные сообщения, несмотря на наличие искажений отдельных символов (в пределах корректирующей способности кода). Не существует кодов конечной длины, позволяющий обнаружить и, тем более, исправить все возможные ошибки. Можно построить лишь коды, обнаруживающие или исправляющие некоторое число ошибок определенного вида (наиболее вероятных или наиболее опасных).

Коды, обнаруживающие и исправляющие ошибки, называются *помехоустойчивыми* или *корректирующими*. Корректирующая способность кода основывается на его избыточности. При отсутствии избыточности, когда любая комбинация символов алфавита канала может представлять передаваемое сообщение, достоверное обнаружение ошибок невозможно: на приемной стороне неизвестно, какое именно сообщение могло передаваться (в противном случае его передача не несла бы информации и была бы лишена смысла). Для того чтобы избыточностью сообщения можно было воспользоваться для обнаружения или исправления ошибок, она должна вноситься по определенным правилам. При рассмотрении корректирующей способности кодов будем учитывать лишь избыточность, вносимую в сообщение при кодировании, сообщения же, поступающие от источника, будем считать безыбыточными. В частном случае, когда алфавиты источника и канала имеют одинаковый объем  $m$ , внесение избыточности означает увеличение числа символов в передаваемой последовательности по отношению к информационной последовательности.

По форме внесения избыточности различают *систематические* и *несистематические* коды. В систематических кодах символы информационной последовательности входят без изменения в передаваемое сообщение, занимая в нем отведенные им позиции, а кодирование сводится к внесению в передаваемое сообщение дополнительных (избыточных) символов, связанных определенной зависимостью с символами информационной последовательности. В несистематических кодах символы информационной последовательности в явном виде в передаваемое сообщение не вводят, а могут быть установлены по известным зависимостям, связывающим их с символами передаваемого сообщения.

По способу кодирования различают *блочные* и *непрерывные* коды. При блочном кодировании информационная последовательность разбивается на блоки символов  $(b_1, b_2, \dots, b_k)$  фиксированной длины  $k$ , каждому из

которых ставится в соответствие определенная комбинация символов алфавита канала ( $\beta_1, \beta_2, \dots, \beta_n$ ), называемая *кодovým* словом. Код при блочном кодировании определяет закон формирования кодового слова, отвечающего данному блоку информационных символов. Для цифровой информации в качестве таких блоков удобно использовать передаваемые числа. В случае систематического блочного кода кодовое слово, отвечающее блоку информационных символов ( $b_1, b_2, \dots, b_k$ ), может быть представлено в виде ( $b_1, b_2, \dots, b_k, \beta_{k+1}, \dots, \beta_n$ ), где  $\beta_{k+1}, \dots, \beta_n$  — избыточные символы. Блочный код, приводящий в соответствие блоку информационных символов длиной  $k$  кодовое слово длиной  $n$ , будем обозначать  $(n, k)$ . Избыточные символы не несут дополнительной информации об источнике сообщения (они однозначно определяются информационными символами  $b_1, b_2, \dots, b_k$ ). Поэтому кодовое слово несет то же количество полезной информации, что и соответствующий блок информационных символов. При безызбыточности входного сообщения блок из  $k$  информационных символов (а значит, и кодовое слово) несет количество информации  $H = k \log_2 m$ , где  $m$  — объем алфавита источника.

Максимальное количество информации, которое может содержать слово из  $n$  символов канала при том же объеме алфавита  $m$ , равно  $H_{\max} = n \log_2 m$ .

Поэтому избыточность кода  $(n, k)$

$$\rho_n = 1 - \frac{H}{H_{\max}} = 1 - \frac{k \log_2 m}{n \log_2 m} = \frac{n - k}{n} = \frac{r}{n}, \quad (8.3)$$

где  $r = n - k$  — число избыточных символов, вносимых при кодировании.

При непрерывном кодировании каждый символ передаваемого сообщения определяется рекуррентными соотношениями, связывающими его с соответствующими символами информационной последовательности:

$$\beta_{ij} = \left[ \sum_{v=-q}^{s-1} c_{vj} b_{i+v} \right] (\text{mod } m), \quad j = 0, 1, \dots, l-1. \quad (8.4)$$

Значение правой части (8.4) определяется "по модулю  $m$ ", что означает, что после проведения обычных вычислений удерживается лишь остаток от деления полученной величины на основание кода  $m$  (например, 9 "по модулю 4" дает 1). Осуществление вычислений в (8.4) "по модулю  $m$ " необходимо для того, чтобы полученные значения символов  $\beta_{ij}$  принадлежали алфавиту канала, т.е. принимали одно из значений  $0, 1, \dots, m-1$ .

Рекуррентное соотношение (8.4) на  $i$ -м шаге кодирования группе символов входной последовательности  $b_{i-q} \dots b_i \dots b_{i+s-1}$  ставит в соответствие  $l$  символов кодовой последовательности  $\beta_{i0}, \dots, \beta_{i,l-1}$ ; на  $(i+1)$ -м шаге группе символов  $b_{i-q+1}, \dots, b_{i+1}, \dots, b_{i+s}$ , смещенной на одну позицию, ста-

вит в соответствие следующие  $l$  символов кодовой последовательности  $\beta_{i+1,0}, \dots, \beta_{i+1,l-1}$  и т.д. Таким образом, каждому символу информационной последовательности на очередном шаге кодирования по формуле (8.4) приводится в соответствие  $l$  символов передаваемой кодовой последовательности, что соответствует избыточности кода  $\rho_n = (l-1)/l$ . Правые части рекуррентных соотношений (8.4) представляют как бы свертку соответствующего участка информационной последовательности, поэтому коды и называются *сверточными*.

Непрерывные (сверточные) коды могут быть как *систематическими*, так и *несистематическими*. Систематические сверточные коды получаются в частном случае, когда для одного из значений индекса  $j$  (например  $j=0$ ) коэффициенты в формуле (8.4) принимают значения

$$C_{v0} = \begin{cases} 0 & \text{при } v \neq 0, \\ 1 & \text{при } v = 0 \end{cases}$$

и соответственно  $\beta_{i0} = b_i$ . В этом случае передаваемая кодовая последовательность имеет вид  $\dots b_i, \beta_{i1}, \dots, \beta_{i,l-1}, b_{i+1}, \beta_{i+1,1}, \dots, \beta_{i+1,l-1}, \dots$  и содержит очередной символ информационной последовательности и  $l-1$  избыточных символов.

В последующих главах проводится более детальное рассмотрение общих свойств и некоторых конкретных классов блочных и непрерывных (сверточных) корректирующих кодов.

## Глава 9 БЛОЧНЫЕ КОРРЕКТИРУЮЩИЕ КОДЫ

### 9.1. Основные характеристики и корректирующие свойства блочных кодов

В блочных корректирующих кодах  $(n, k)$  любому блоку информационных символов  $(b_1, b_2, \dots, b_k)$  длиной  $k$  ставится в соответствие кодовое слово  $b = (\beta_0, \beta_1, \dots, \beta_{n-1})$ , содержащее  $n$  символов алфавита канала. Будем считать, что символы информационной последовательности и кодового слова принадлежат одному и тому же алфавиту объема  $m$  и могут принимать значения  $0, 1, \dots, m-1$ . Из общего количества кодовых слов данной длительности, которые могут быть образованы с использованием алфавита канала, равного  $m^n$ , для передачи используется лишь  $m^k$  так называемых разрешенных кодовых слов, что позволяет обнаруживать ошибки, если только в результате искажений переданное кодовое слово не трансформировалось в другое разрешенное кодовое слово.

В случае, когда искажения отдельных символов кодового слова могут рассматриваться как случайные независимые события, вероятность соответствующих искажений кодового слова уменьшается с увеличением кратности ошибок (числа искаженных символов в слове). При этом вероятность трансформации одного кодового слова в другое тем меньше, чем большим числом символов они различаются. Это оправдывает использование для отбора множества  $B_p$  разрешенных кодовых слов метрики Хэмминга, согласно которой расстояние  $d_{ij}$  между кодовыми словами  $b^{(i)}$  и  $b^{(j)}$  определяется числом позиций с несовпадающими символами. Например, коды 100111 и 110110, отличающиеся символами в двух разрядах, имеют расстояние по Хэммингу, равное двум.

Минимальное из расстояний  $d_{ij}$  между разрешенными кодовыми словами

$$d_{\min} = \min_{i,j} \{d_{ij}\}, \quad b^{(i)}, b^{(j)} \in B_p,$$

называется *кодovým расстоянием*. Расстояние между переданным ( $b$ ) и опознанным ( $b'$ ) кодовыми словами при использовании метрики Хэмминга определяется числом искаженных символов.

Для исправления ошибок при случайном независимом характере искажений отдельных символов рационально использовать следующее правило декодирования: из всех разрешенных кодовых слов  $b^{(i)} \in B_p$  в качестве оценки принятого слова  $b'$  выбирается то, которое минимизирует расстояние  $d(b', b^{(i)})$ . Минимальному расстоянию  $d(b', b^{(i)})$  соответствует минимум искажений переданного слова  $b^{(i)}$ . Поэтому при приеме слова  $b'$  гипотеза о том, что передавалось слово  $b^{(i)}$ , является наиболее вероятной. Таким образом, сформулированное правило декодирования соответствует критерию максимального правдоподобия (см. § 3.2).

В случае, когда распознавание отдельных символов ведется без учета корреляционных связей между ними (так называемый посимвольный прием), а распознавание кодовых слов (декодирование) производится без учета корреляционных связей между отдельными словами, составляющими сообщение, это правило может считаться оптимальным. Однако при реализации приема сообщения в целом, когда принятие решения производится с учетом не только условных вероятностей искажений отдельных разрешенных кодовых слов, приводящих к их трансформации в принятое слово  $b'$ , но и априорных вероятностей тех или иных реализаций сообщения в целом, оптимальное правило декодирования должно основываться на критерии максимума апостериорной вероятности сообщения. Естественно, что прием в целом, когда при декодировании учитываются все статистические свойства источника сообщений, обеспечивает более высокую достоверность. Например, при чтении неразборчивой рукописи, используя для распознавания отдельных слов контекст, человек эвристически реализует "прием в целом". В связи с технической трудностью реализации приема в целом в практике связи в настоящее время наиболее часто используется посимвольный прием и декодирование по критерию максимального правдоподобия. Правда, за последнее время все большее внимание привлекают так называемые методы последовательного декодирования, позволяющие реализовать прием в целом для участков сообщения заданной длительности (см. § 10.3). В дальнейшем всюду (кроме специально отмеченных случаев) будет подразумеваться посимвольный прием.

В канале без стирания, когда каждый символ на выходе обязательно идентифицируется с одним из символов алфавита, максимальное число ошибок  $\mu_{\max}$ , при котором факт искажения кодового слова достоверно обнаруживается, определяется равенством

$$u_{\max} = d_{\min} - 1. \quad (9.1)$$

При числе ошибок  $u \leq d_{\min} - 1$  искаженное кодовое слово не может совпасть ни с одним из разрешенных слов и факт искажения обнаруживается всегда. Могут быть случаи обнаружения факта наличия ошибок и при  $u > d_{\min} - 1$ , но не всегда (а значит, недостоверно). Максимальное число достоверно исправляемых ошибок  $t_{\max}$  определяется неравенством

$$t_{\max} \leq (d_{\min} - 1)/2 \quad (9.2)$$

При выполнении условия (9.2) ближайшим по метрике Хэмминга к принятому обязательно будет переданное кодовое слово и сформулированное правило декодирования обеспечит правильное исправление ошибок.

Следует иметь в виду, что исправление ошибок приводит к повышению вероятности имитации противником разрешенных кодовых слов при организации помех (понижению имитостойкости канала). Понижение имитостойкости будет тем большим, чем выше корректирующая способность исправляющего ошибки кода. При полном использовании корректирующей способности такого кода, когда любое принятое кодовое слово идентифицируется с одним из разрешенных слов, для создания имитационной помехи достаточно воспроизводить в любой последовательности символы алфавита канала без учета структуры кода. Поэтому не всегда имеет смысл полностью реализовывать корректирующую способность кода при исправлении ошибок. В высоконадежных каналах с относительно малой вероятностью искажения отдельных символов для повышения имитостойкости целесообразно искусственно ограничивать число исправляемых ошибок  $t$  значением, меньшим  $(d - 1)/2$ , отождествляя принятые кодовые слова с одним из разрешенных только в случае, если расстояние между ними не более  $t$ . В этом случае, исправляя до  $t$  ошибок, можно достоверно обнаружить факт искажения кодового слова при числе ошибок, большем  $t$ , но не превышающем величину  $u_{\max}$ , удовлетворяющую условию:

$$u_{\max} = d_{\min} - t - 1; \quad t < (d_{\min} - 1)/2. \quad (9.3)$$

При большем числе ошибок принятое слово может оказаться на расстоянии, не превосходящем  $t$  от другого (непередаваемого) слова, и будет с ним идентифицировано. Таким образом, в рассматриваемом случае при

числе ошибок  $u \leq t$  они достоверно исправляются, при числе ошибок  $t < u \leq d - t - 1$  достоверно обнаруживается факт искажения переданного кодового слова, а при  $t > u > d - t - 1$  может иметь место неверное исправление принятого слова.

В каналах со стиранием достоверное восстановление переданного кодового слова при отсутствии трансформированных символов обеспечивается при числе стертых символов  $s$ , удовлетворяющем условию:

$$s \leq d_{\min} - 1. \quad (9.4)$$

При этом для идентификации принятого кодового слова с переданным достаточно ввести в рассмотрение "укороченный" код, исключив как в принятом, так и во всех разрешенных кодовых словах позиции, содержащие стертые символы в принятом слове. Переданное кодовое слово соответствует тому разрешенному слову, которое совпадает с принятым в "укороченном" коде. Кодовое расстояние для "укороченного" кода может уменьшиться по сравнению с исходным на число исключенных позиций и будет не менее  $d_{\min} - s$ . Поэтому при наличии  $s$  стертых символов имеются дополнительные возможности обнаружения и исправления ошибок (помимо восстановления стертых символов), определяемые: максимальным числом достоверно обнаруживаемых ошибок, в соответствии с (9.1) равным

$$u_{\max} = d_{\min} - s - 1; \quad (9.5)$$

максимальным числом достоверно исправляемых ошибок, в соответствии с (9.2) равным

$$t_{\max} \leq (d_{\min} - s - 1)/2; \quad (9.6)$$

максимальным числом достоверно обнаруживаемых ошибок при исправлении  $t$  ошибок, в соответствии с (9.3) равным

$$u_{\max} = d_{\min} - s - t - 1; \quad t < (d_{\min} - s - 1)/2. \quad (9.7)$$

Кодовое расстояние, определяющее корректирующую способность кода, зависит от избыточности последнего. В кодах без избыточности, когда для передачи используются все возможные кодовые комбинации (так

называемые простые или примитивные коды), ближайшие разрешенные кодовые слова отличаются всего одним символом ( $d_{\min} = 1$ ) и такие коды не обладают корректирующей способностью. Кодовое расстояние будет тем большим, чем большей будет избыточность кода и чем равномернее будут расстояния между разрешенными кодовыми словами. Коды, у которых заданная корректирующая способность достигается при минимальной избыточности  $r_{\min}$ , называются *оптимальными*. Коды, которые исправляют все ошибки заданного вида, и при этом множества  $B_p$  разрешенных и  $B_n$  искаженных заданными ошибками кодовых слов исчерпывают все возможные кодовые слова (при заданных длине и основании кода), называются *совершенными*. Оптимальный код может не быть совершенным.

Рядом авторов получены соотношения, определяющие связь между избыточностью и кодовым расстоянием для оптимальных кодов. Так, граница Плоткина дает верхнюю границу кодового расстояния  $d_{\min}$  при заданных основании кода  $m$ , числе элементов кода  $n$  и числе информационных символов  $k$  [26]:

$$d_{\min} (m^k - 1) \leq nm^{k-1} (m - 1). \quad (9.8)$$

Граница Варламова—Гилберта дает нижнюю границу для числа избыточных символов  $r$ , необходимого для обеспечения кодового расстояния  $d_{\min}$  [26]:

$$m^r \geq 1 + \sum_{i=1}^{d_{\min}-2} C_{n-1}^i (m-1)^i, \quad (9.9)$$

где  $C_{n-1}^i$  — число сочетаний из  $n-1$  элемента по  $i$  элементов. В частности, в случае двоичного кода ( $m=2$ ) неравенство (9.9) принимает вид

$$2^r \geq 1 + \sum_{i=1}^{d_{\min}-2} C_{n-1}^i, \quad (9.10)$$

Если случайные независимые искажения отдельных символов кодового слова характеризуются вероятностью  $P_e$ , то вероятность неправильного опознавания слова при простом коде, содержащем  $n$  символов,

$$P_{\text{сл}}^{(\text{прост})} = 1 - (1 - P_e)^n. \quad (9.11)$$

Вероятность ошибки при использовании исправляющего корректирующего кода с кодовым расстоянием  $d_{\min}$  не превышает вероятности появ-

ления ошибок с кратностью, больше  $t$ , и в соответствии с (2.7) удовлетворяет условию

$$P_{\text{сл}}^{(\text{исп})} \leq 1 - \sum_{i=0}^t C_n^i P_e^i (1 - P_e)^{n-i}, \quad t \leq (d_{\min} - 1)/2. \quad (9.12)$$

Вероятность  $P_{\text{сл}}^{(\text{исп})}$ , как видно из сравнения (9.11) и (9.12), меньше вероятности ошибки для простого кода на величину

$$\sum_{i=1}^t C_n^i P_e^i (1 - P_e)^{n-i}.$$

Вероятность необнаружения ошибок при использовании обнаруживающего корректирующего кода с тем же кодовым расстоянием определяется неравенством:

$$P_{\text{сл}}^{(\text{обн})} \leq 1 - \sum_{i=0}^{d_{\min}-1} C_n^i P_e^i (1 - P_e)^{n-i}, \quad (9.13)$$

т.е. еще меньше.

Из проведенного рассмотрения общих свойств блочных корректирующих кодов следует, что и по своему характеру, и по результатам задачи оптимального помехоустойчивого кодирования последовательностей символов ограниченной длительности принципиально отличается от задачи оптимального кодирования по Шеннону бесконечно длинных последовательностей. Прежде всего такие понятия, как обнаружение факта наличия ошибок, число обнаруженных и исправляемых ошибок имеют смысл только при рассмотрении кодов ограниченной длины. При рассмотрении бесконечно длинных кодовых слов эти понятия лишены смысла: в типичных последовательностях бесконечной длительности при заданной ненадежности канала наличие ошибок обязательно, а число ошибок каждого вида определено статистическими характеристиками канала. Что касается результатов оптимального кодирования, то при кодировании бесконечно длинных последовательностей символов по Шеннону обеспечивается сколь угодно малая вероятность ошибки при минимальной избыточности кода, определяемой соотношением (7.35):

$$p \approx H(B|B') / \log_2 m,$$

где  $H(B|B')$  — ненадежность канала;  $m$  — объем алфавита канала. При кодировании же блоков символов ограниченной длительности вероятность ошибочного опознавания сообщения, как следует из (9.12), хотя и может быть сделана весьма малой, но остается конечной, избыточность же существенно больше минимально необходимой по Шеннону.

В принципе блочный корректирующий код может быть реализован в виде таблицы разрешенных кодовых слов, составленной любым способом (например, подбором). Это так называемые *случайные корректирующие коды*. Обнаружение и исправление ошибок при этом сводится к сравнению принятого кодового слова со всеми разрешенными кодовыми словами, хранящимися в памяти; несовпадение принятого слова ни с одним из разрешенных свидетельствует о наличии ошибок. Исправление ошибок сводится к идентификации принятого слова с ближайшим к нему по метрике Хэмминга разрешенным кодовым словом.

Однако такая процедура громоздка, требует большого объема памяти и практически может применяться лишь в малоинформативных каналах, использующих небольшое число различных кодовых слов. Для высокоинформативных каналов случайные корректирующие коды неприемлемы и применение нашли только так называемые *регулярные корректирующие коды*. В таких кодах процедуры отбора разрешенных кодовых слов, обнаружения и исправления ошибок определяются регулярными правилами кодирования и декодирования и в памяти хранятся только эти правила. Для построения регулярных блочных кодов успешно используется аппарат линейной алгебры. При этом блоки информационных символов и кодовые слова отождествляются с соответствующими математическими образами (векторами, полиномами) и операции кодирования и декодирования сводятся к их алгебраическим преобразованиям по определенным законам.

Рассмотрим некоторые классы регулярных блочных корректирующих кодов.

## 9.2. Линейные корректирующие коды. Коды Хэмминга

Математическая теория помехоустойчивого кодирования может, в частности, базироваться на теории линейных векторных пространств. При этом общая теория строится для случаев, когда основание кода  $m$  — простое число, т.е. делится лишь само на себя и на единицу, а число разрядов  $n$  фиксировано (равномерный код).

При использовании теории линейных векторных пространств кодовое слово  $b$  длиной  $n$  рассматривается как вектор в линейном  $n$ -мерном

пространстве кодовых слов  $R_n$ , а символы  $b_0, b_1, \dots, b_{n-1}$  — как его компоненты.

*Линейным* называется блочный равномерный код, для которого любая линейная композиция кодовых слов (двух или более), принадлежащих множеству разрешенных слов  $B_p$ , дает разрешенное кодовое слово, т.е. для любых векторов  $b^{(i)}, b^{(j)} \in B_p$  и любых значений символов  $\lambda_i, \lambda_j$ , принадлежащих алфавиту канала, справедливо условие

$$b^{(k)} = \lambda_i b^{(i)} + \lambda_j b^{(j)} \in B_p. \quad (9.14)$$

Следует уточнить операции умножения и суммирования символов при линейных преобразованиях кодовых слов. Для того чтобы компоненты результирующего вектора  $b^{(k)}$  состояли только из символов канала  $0, 1, \dots, m-1$  (а это необходимо для того, чтобы  $b^{(k)} \in B_p$ ), операции умножения и сложения компонент векторов  $b^{(i)}$  и  $b^{(j)}$  при линейном преобразовании (9.14) определяются по модулю  $m$ . Например, при операциях с троичными символами  $0, 1, 2$  получим:

$$(2+1) \pmod 3 = 0; \quad (2+2) \pmod 3 = 1; \quad (2 \cdot 2) \pmod 3 = 1.$$

Осуществление операций умножения и сложения символов кода по модулю  $m$  придает им свойство замкнутости: в результате этих операций получаются символы, принадлежащие исходному множеству. Операции умножения и сложения по модулю  $m$  сохраняют свойства ассоциативности, дистрибутивности и коммутативности, присущие обычным операциям умножения и сложения. В линейной алгебре множество, удовлетворяющее условию замкнутости по отношению к двум заданным операциям, обладающим свойствами ассоциативности, дистрибутивности и коммутативности, называется полем Галуа и обозначается  $GF(m)$ . Тот факт, что по отношению к используемым в преобразовании (9.14) операциям символы кода образуют поле Галуа, позволяет говорить (в терминах линейной алгебры), что линейное пространство векторов кодовых слов  $R_n$  строится над полем Галуа  $GF(m)$ .

В пространстве  $R_n$  кодовых слов вводится также скалярное произведение векторов  $a$  и  $b$ :

$$(ab) = \left[ \sum_{v=0}^{n-1} a_v b_v \right] \pmod m. \quad (9.15)$$

Векторы (кодовые слова), для которых  $(ab) = 0$ , называются *ортогональными*.

Представление передаваемых и опознанных кодовых слов в линейном векторном пространстве  $R_n$  позволяет математически описать искажения кодовых слов вектором  $e \in R_n$ , удовлетворяющим условию

$$\mathbf{b}' = \mathbf{b} + \mathbf{e}, \quad (9.16)$$

где  $\mathbf{b}$  и  $\mathbf{b}'$  — вектора передаваемого и опознанного кодовых слов.

Для задания линейного кода можно воспользоваться одним из следующих его свойств:

1. Код, содержащий  $M = m^k$  кодовых слов в разрешенном наборе, может быть задан  $k$  линейно-независимыми векторами  $\mathbf{b}^{(i)}$ ,  $i = 1, 2, \dots, k$ , составляющими базисную матрицу  $L$  размером  $k \times n$ <sup>\*)</sup>:

$$L = \begin{bmatrix} \mathbf{b}^{(1)} \\ \dots \\ \mathbf{b}^{(k)} \end{bmatrix} = \begin{bmatrix} \beta_0^{(1)} & \beta_1^{(1)} & \dots & \beta_{n-1}^{(1)} \\ \dots & \dots & \dots & \dots \\ \beta_0^{(k)} & \beta_1^{(k)} & \dots & \beta_{n-1}^{(k)} \end{bmatrix}. \quad (9.17)$$

Множество разрешенных кодовых слов, определяемое базисной матрицей  $L$ , обозначим  $B^L$ . Любое кодовое слово из разрешенного набора  $B^L$ , определяемого матрицей (9.17), может быть представлено в виде линейной композиции входящих в нее векторов:

$$\mathbf{b}^{(i)} = \sum_{i=1}^k \lambda_i \mathbf{b}^{(i)}, \quad \lambda_i \in GF(m). \quad (9.18)$$

Нетрудно видеть, что общее число разрешенных кодовых слов, образуемых по закону (9.18), равно числу различных комбинаций из  $k$  коэффициентов  $\lambda_i$ , принимающих одно из значений  $0, 1, \dots, m-1$  т.е. закон (9.18) обеспечивает образование заданного числа  $M = m^k$  различных кодовых слов длиной  $n$ . Таким образом, базисная матрица размером  $k \times n$  задает линейный код  $(n, k)$ .

Простой код является линейным кодом  $(n, n)$ , базисная матрица которого включает полный набор  $n$ -мерных линейно-независимых векторов (число линейно-независимых векторов в  $n$ -мерном пространстве не может превосходить  $n$ ).

2. Множеству разрешенных кодовых слов  $B^L$ , определяемому матрицей  $L$ , отвечает некоторое множество  $B^H$  кодовых слов, ортогональных кодовым словам разрешенного набора: если  $\mathbf{b}^{(i)} \in B^L$ , а  $\mathbf{b}^{(j)} \in B^H$ , то  $(\mathbf{b}^{(i)} \mathbf{b}^{(j)}) = 0$ . Множество  $B^H$  может быть задано  $n-k$  линейно-независимыми векторами, составляющими матрицу  $H$  размером  $(n-k) \times n$ :

$$H = \begin{bmatrix} \mathbf{b}^{(k+1)} \\ \vdots \\ \mathbf{b}^{(n)} \end{bmatrix} = \begin{bmatrix} \beta_0^{(k+1)} & \beta_1^{(k+1)} & \dots & \beta_{n-1}^{(k+1)} \\ \dots & \dots & \dots & \dots \\ \beta_0^{(n)} & \beta_1^{(n)} & \dots & \beta_{n-1}^{(n)} \end{bmatrix}.$$

Матрица  $H$  называется *проверочной*. Множество  $B^H$  содержит  $m^{(n-k)}$  кодовых слов, отвечающих всем возможным линейным композициям векторов, составляющих матрицу  $H$ . Свойства множеств  $B^L$  и  $B^H$  взаимны: множество  $B^H$  также представляет линейный код, для которого матрица  $H$  является базисной, а матрица  $L$  — проверочной.

Проверочная матрица  $H$  так же, как и базисная матрица  $L$ , полностью определяет линейный код. Она позволяет представить линейный код  $(n, k)$  в систематической форме:

$$\mathbf{b}^{(i)} = (b_0^{(i)}, b_1^{(i)}, \dots, b_{k-1}^{(i)}, \beta_k^{(i)}, \dots, \beta_{n-1}^{(i)}). \quad (9.19)$$

где первые  $k$  символов  $b_0^{(i)}, b_1^{(i)}, \dots, b_{k-1}^{(i)}$  — информационные, а последние  $n-k$  символов  $\beta_k^{(i)}, \beta_{k+1}^{(i)}, \dots, \beta_{n-1}^{(i)}$  — избыточные, обеспечивающие возможность обнаружения и исправления ошибок (их называют проверочными символами).

Проверочные символы разрешенного кодового слова  $\mathbf{b}^{(i)}$  при заданных информационных символах  $b_0^{(i)}, b_1^{(i)}, \dots, b_{k-1}^{(i)}$  могут быть определены из  $n-k$  линейных уравнений, определяемых условиями ортогональности разрешенного кодового слова кодовым словам проверочной матрицы  $H$ :

$$S_v = (\mathbf{b}^{(i)} \mathbf{b}^{(k+v)}) = 0, \quad v = 1, 2, \dots, n-k. \quad (9.20)$$

Уравнения (9.20) можно переписать в виде:

$$\sum_{j=k}^{n-1} b_j^{(k+v)} \beta_j^{(i)} = \sum_{j=0}^{k-1} b_j^{(i)} b_j^{(k+v)}, \quad v = 1, 2, \dots, n-k. \quad (9.21)$$

Разрешая систему  $n-k$  уравнений (9.21) относительно  $n-k$  проверочных символов  $\beta_j^{(i)}$ , получим для них линейные выражения вида

$$\beta_j^{(i)} = \sum_{l=0}^{k-1} c_{lj} b_l^{(i)}, \quad j = k, k+1, \dots, n-1. \quad (9.22)$$

Формула (9.22) позволяет определить  $n-k$  проверочных символов  $\beta_k, \beta_{k+1}, \dots, \beta_{n-1}$ , разрешенного кодового слова по известным  $k$  информационным символам (коэффициенты  $c_{lj}$  определяющие закон формирования проверочных символов, выражаются через элементы  $b_j^{(k+v)}$  проверочной матрицы  $H$ ).

<sup>\*)</sup> Краткие сведения о матрицах см. в приложении.

Формула (9.22) является весьма удобной формой задания регулярного линейного кода  $(n, k)$ , широко используемой в практике связи. При этом для образования разрешенных кодовых слов по известным информационным символам достаточно хранить в памяти  $k(n-k)$  значений коэффициентов  $c_{ij}$ , входящих в (9.22). Поскольку при задании кода с основанием  $m$  для каждого из этих коэффициентов может быть выбрано одно из  $m$  значений, то всегда можно выбрать  $m^{k(n-k)}$  вариантов коэффициентов в (9.22) и построить  $m^{k(n-k)}$  различных линейных кодов  $(n, k)$ . Задача заключается в том, чтобы выбрать оптимальный вариант кода.

Критерием наличия ошибок в принятом слове  $b'$  является нарушение условий его ортогональности векторам  $b^{(k+v)}$  проверочной матрицы. Результаты проверки соблюдения условий ортогональности дают  $n-k$  значений скалярных произведений.

$$S_v = (b'b^{(k+v)}) = \sum_{j=0}^{n-1} b'_j b_j^{(k+v)}, \quad v = 1, 2, \dots, n-k, \quad (9.23)$$

которые могут рассматриваться в качестве компонент некоторого вектора  $S$ , называемого *синдромом проверки*. Для неискаженных кодовых слов разрешенного набора, удовлетворяющих условиям ортогональности векторам  $b^{(k+v)}$ ,

$$S_v = 0, \quad v = 1, 2, \dots, n-k, \quad (9.24)$$

т.е. синдром  $S$  представляет нулевой вектор, все компоненты которого равны нулю. При наличии искажений синдром позволяет установить их характер и произвести исправление некоторых видов ошибок.

Для двоичных кодов скалярные произведения (9.23) вырождаются в так называемые контрольные суммы, представляющие суммы по модулю 2 символов двоичного кода  $b'$ , отвечающих ненулевым позициям  $v$ -й строки проверочной матрицы. При этом равенство нулю синдрома обеспечивается в случае четного числа единиц в коде  $b'$  на позициях, входящих в каждую из контрольных сумм. Поэтому для двоичного кода проверку на ортогональность (9.23) называют *проверкой на четность*.

В качестве примера рассмотрим весьма распространенный класс линейных кодов — двоичные коды Хэмминга. В двоичном коде Хэмминга для обнаружения и исправления ошибок определяются контрольные суммы по модулю 2, каждая из которых включает некоторое число информационных символов и один из проверочных символов. Обнаруживающий ошибки код Хэмминга содержит лишь один проверочный символ (в младшем разряде), который принимается равным сумме по модулю 2 всех информационных символов

$$\beta_{n-1} = \sum_{j=0}^{n-2} b_j \pmod{2}. \quad (9.25)$$

Формула (9.25) является частным случаем формулы (9.22) при  $m=2$ ,  $k=n-1$  и  $c_{j\mu} = 1$ ,  $\mu = 0, 1, \dots, n-2$ . Единственной контрольной суммой этого кода является сумма по модулю 2 всех символов, включая проверочный:

$$\left[ \sum_{j=0}^{n-2} b_j + \beta_{n-1} \right] \pmod{2} = 0. \quad (9.26)$$

Формула (9.26) является частным случаем формулы (9.21), когда  $k=n-1$ , а все составляющие единственного вектора, входящего в проверочную матрицу, равны единице:  $b^{(n)}_j = 1$ ,  $j = 0, 1, \dots, n-1$ . Очевидно, что при отсутствии искажений в канале контрольная сумма (9.26) для опознанного кодового слова равна нулю, а при любом одиночном искажении (или нечетном числе искаженных символов) равна единице. Отличие от нуля контрольной суммы обнаруживает факт (но не место) одиночного или нечетного числа искажений символов принятого кодового слова. При четном числе искажений последние не обнаруживаются. Таким образом, данный код достоверно обнаруживает лишь одиночные ошибки. Кратные ошибки обнаруживаются при нечетном их числе. Обнаруживающий ошибки код Хэмминга обладает кодовым расстоянием  $d_{\min} = 2$ , так как ближайшее к передаваемому разрешенное кодовое слово, отличающееся одним информационным символом, обязательно отличается и значением проверочного символа. Расстояние  $d_{\min} = 2$  до ближайшего разрешенного кодового слова имеет место для любого передаваемого слова, что соответствует максимально возможному при данном кодовом расстоянии набору разрешенных кодовых слов. Поэтому код Хэмминга, обнаруживающий одиночные ошибки, является оптимальным. Одновременно он является и совершенным кодом для обнаружения одиночных ошибок, поскольку все возможные кодовые слова либо принадлежат разрешенному набору, либо соответствуют случаю искажения одного символа разрешенного слова.

В исправляющем коде Хэмминга двоичное число, составленное из контрольных сумм (синдром кода) должно определять порядковый номер искаженного разряда кодового слова. Для этого контрольные суммы составляются так, чтобы в сумму  $S_j$  входили символы кодового слова, имеющие единицу в  $j$ -м разряде в двоичной записи их порядкового номера. Нетрудно видеть, что при этом символы, порядковый номер которых в двоичной записи выражается единицей  $j$ -го разряда, т.е. равен  $2^{j-1}$ ,  $j = 1, 2, 3, \dots$ , войдут только в одну контрольную сумму  $S_j$ . Именно на этих позициях (1-й, 2-й, 4-й, 8-й и т.д.) в коде Хэмминга и располагаются проверочные символы  $\beta_1, \beta_2, \beta_4, \beta_8, \dots$ . Соответствующим их выбором все



контрольные суммы обращаются в нуль. При отсутствии искажений все контрольные суммы останутся равными нулю. Одиночная ошибка, например искажение  $l$ -го разряда кодового слова, приведет к обращению в единицу контрольных сумм, номера которых соответствуют номерам разрядов двоичной записи числа  $l$ , содержащих единицы. Поэтому, записывая значения контрольных сумм в порядке убывания их номеров, т.е. записывая синдром  $\dots S_4 S_3 S_2 S_1$ , получаем в двоичной записи номер  $l$  искаженного разряда.

Для наглядности двоичная запись номеров разрядов кодового слова сведена в табл. 9.1 (таблица приведена для 15-разрядных кодов).

Таблица 9.1

| Порядковый номер разряда кода в десятичной записи | Значение разрядов двоичной записи порядкового номера разрядов кода |   |   |   |
|---|--|---|---|---|
|   | 1<br>(младший)   | 2 | 3 | 4 |
| 1 (младший)                                       | 1  | 0 | 0 | 0 |
| 2   | 0  | 1 | 0 | 0 |
| 3   | 1  | 1 | 0 | 0 |
| 4   | 0  | 0 | 1 | 0 |
| 5   | 1  | 0 | 1 | 0 |
| 6   | 0  | 1 | 1 | 0 |
| 7   | 1  | 1 | 1 | 0 |
| 8   | 0  | 0 | 0 | 1 |
| 9   | 1  | 0 | 0 | 1 |
| 10  | 0  | 1 | 0 | 1 |
| 11  | 1  | 1 | 0 | 1 |
| 12  | 0  | 0 | 1 | 1 |
| 13  | 1  | 0 | 1 | 1 |
| 14  | 0  | 1 | 1 | 1 |
| 15  | 1  | 1 | 1 | 1 |

Для 15-разрядного кода в соответствии с табл. 9.1 контрольные суммы имеют вид:

$$\begin{aligned} S_1 &= \beta_1 \oplus b_3 \oplus b_5 \oplus b_7 \oplus b_9 \oplus b_{11} \oplus b_{13} \oplus b_{15}, \\ S_2 &= \beta_2 \oplus b_3 \oplus b_6 \oplus b_7 \oplus b_{10} \oplus b_{11} \oplus b_{14} \oplus b_{15}, \\ S_3 &= \beta_4 \oplus b_5 \oplus b_6 \oplus b_7 \oplus b_{12} \oplus b_{13} \oplus b_{14} \oplus b_{15}, \\ S_4 &= \beta_8 \oplus b_9 \oplus b_{10} \oplus b_{11} \oplus b_{12} \oplus b_{13} \oplus b_{14} \oplus b_{15}, \end{aligned} \quad (9.27)$$

где  $b_i$  — значения информационных символов кода;  $\beta_1, \beta_2, \beta_4, \beta_8$  — значения проверочных символов;  $\oplus$  — знак суммирования по модулю два.

Значения проверочных символов  $\beta_i$ , обращающие в нуль контрольные суммы (9.27), равны

$$\begin{aligned} \beta_1 &= b_3 \oplus b_5 \oplus b_7 \oplus b_9 \oplus b_{11} \oplus b_{13} \oplus b_{15}, \\ \beta_2 &= b_3 \oplus b_6 \oplus b_7 \oplus b_{10} \oplus b_{11} \oplus b_{14} \oplus b_{15}, \\ \beta_4 &= b_5 \oplus b_6 \oplus b_7 \oplus b_{12} \oplus b_{13} \oplus b_{14} \oplus b_{15}, \\ \beta_8 &= b_9 \oplus b_{10} \oplus b_{11} \oplus b_{12} \oplus b_{13} \oplus b_{14} \oplus b_{15}. \end{aligned} \quad (9.28)$$

Процедура формирования исправляющего кода Хэмминга включает, таким образом, следующие операции:

1) расположение информационных символов в порядке возрастания разрядов с сохранением незанятыми разрядов с номерами  $2^{i-1}$ , т.е. 1-го, 2-го, 4-го, 8-го и т.д.;

2) вычисление проверочных символов, занимающих разряды с номерами  $2^{i-1}$ , по формулам (9.28).

Процедура определения передаваемого информационного кода при использовании исправляющего кода Хэмминга включает:

1) вычисление контрольных сумм по формулам (9.27);

2) определение порядкового номера искаженного символа, если не все контрольные суммы равны нулю, и исправление его на другой двоичный символ;

3) выделение информационного кода исключением разрядов с номерами  $2^{i-1}$ ,  $i = 1, 2, \dots$ .

Например, если информационный код имеет вид 10001101 (всюду будем пользоваться записью двоичных чисел в порядке убывания номеров разрядов слева направо), то его символы располагаются в коде Хэмминга в следующей последовательности:  $100\beta_8 110\beta_4 1\beta_2 \beta_1$ . Определив по формулам (9.28) значения проверочных символов  $\beta_1, \beta_2, \beta_3, \beta_4$  получим следующий код Хэмминга 10011100101 (проверочные символы выделены жирным шрифтом). Если, например, при приеме 5-й разряд этого кода ошибочно будет принят как единица, то, вычисляя контрольные суммы, получим  $S_1 = 1$ ;  $S_2 = 0$ ;  $S_3 = 1$ ;  $S_4 = 0$ . Значение синдрома, определяющего номер искаженного разряда в двоичной записи, равно  $S_4 S_3 S_2 S_1 = 0101$ , т.е. для исправления ошибки требуется изменить символ в 5-м разряде.

Правильное исправление кодового слова при использовании кода Хэмминга имеет место для одиночных ошибок, что находится в полном соответствии с кодовым расстоянием, которое в данном случае равно трем. Действительно, если в кодовом слове изменить информационный символ, двоичная запись порядкового номера которого содержит минимальное число единиц, а именно, две единицы (одну единицу в двоичной записи имеет лишь порядковый номер контрольного символа), то одновременно изменятся два контрольных символа, входящие в те же контрольные суммы. Таким образом, ближайшее к каждому передаваемому слову разрешенное кодовое слово отличается тремя символами. Поэтому исправляющий код Хэмминга, обеспечивая максимальный набор разрешенных кодовых слов при  $d_{\min} = 3$  является оптимальным. Заметим, что

этот код, если его использовать не в качестве исправляющего, а в качестве обнаруживающего ошибки, как и положено при  $d_{\min} = 3$ , достоверно обнаруживает и 2 ошибки. Действительно, номера двух искаженных символов в двоичной записи отличаются хотя бы одним разрядом, а это значит, что хотя бы в одну контрольную сумму войдет только один из искаженных символов и она изменится.

Нетрудно видеть, что для образования исправляющего кода Хэмминга уже при одном информационном символе требуются два проверочных символа (в двух младших разрядах); при числе информационных символов от 2 до 4 — три проверочных символа (в 1, 2 4-м разрядах); при числе информационных символов от 5 до 11 — четыре проверочных символа (в 1, 2, 4 и 8-м разрядах) и т.д.

Помимо рассмотренных обнаруживающего и исправляющего кодов Хэмминга, можно использовать комбинированный код, исправляющий одиночные ошибки и достоверно обнаруживающий двойные ошибки. Для этого исправляющий код Хэмминга дополняется еще одним младшим разрядом, равным сумме (по модулю два) всех его разрядов. В этом случае при приеме составляются контрольные суммы исправляющего кода без учета младшего разряда и, кроме того, сумма всех разрядов принятого кода (по модулю два). Если часть контрольных сумм равна единице, то исправление принятого слова по приведенным правилам производится только при условии, что сумма всех символов также равна единице. Если же сумма всех символов равна нулю, что свидетельствует о наличии двух или четного числа ошибок, то кодовое слово не исправляется (ибо правильно исправляются только одиночные ошибки), а лишь устанавливается факт наличия двух (или четного числа) ошибок. Таким образом, при числе ошибок не более двух рассмотренный код *достоверно* исправляет одиночные и обнаруживает двукратные ошибки (при большей кратности возможны пропуск или неверное исправление ошибок). Нетрудно показать, что кодовое расстояние для этого кода равно четырем. Действительно, как было показано, без учета младшего разряда разрешенные кодовые слова отличаются не менее чем тремя символами. Но кодовые слова, отличающиеся в остальных разрядах тремя символами, обязательно (по условию четности) отличаются и младшими разрядами, т.е. разрешенные кодовые слова в целом отличаются не менее чем четырьмя символами. Таким образом, и в этом случае имеет место предельная корректирующая способность (исправление одной и обнаружение двух ошибок), соответствующая кодовому расстоянию  $d_{\min} = 4$ . Код обеспечивает максимальный набор разрешенных кодовых слов при  $d_{\min} = 4$  и является оптимальным.

### 9.3. Циклические коды

Формирование циклических кодов основано на отображении кодовых слов элементами алгебры полиномов. Это позволяет помимо операций суммирования кодовых слов и умножения их на символ  $\lambda$ , использовавшихся в алгебре векторов, дополнительно ввести операции умножения и деления кодовых слов и получить за счет этого дополнительные признаки разрешенных кодовых слов. Кодовое слово, длиной  $n$ , описывается полиномом степени  $n - 1$ , коэффициентами которого являются символы кодового слова:

$$b = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta^{n-1} x^{n-1}. \quad (9.29)$$

Для того чтобы в результате операций, используемых для формирования циклических кодов, получались полиномы, описывающие данный код, при всех операциях над полиномами должны выполняться два условия:

1) действия над коэффициентами полиномов (сложение, вычитание, умножение на символ  $\lambda$ ) должны производиться по модулю  $m$ ; 2) умножение полиномов должно производиться по модулю  $x^n - 1$ , т.е. за результат умножения должен приниматься остаток от деления обычного произведения полиномов на двучлен  $x^n - 1$ . Первое условие необходимо для того, чтобы коэффициенты получаемых полиномов принадлежали алфавиту канала, а второе — чтобы степень этих полиномов соответствовала длине данного кода  $n$ .

В дальнейшем во всех случаях, кроме специально оговариваемых, выполнение этих условий будет подразумеваться. Произведение полиномов по модулю  $x^n - 1$  будем обозначать символом  $\otimes$ , сумму коэффициентов полиномов по модулю  $m$  — символом  $\oplus_m$ , а разность — символом  $\ominus$ .

Циклический код обычно задается *порождающим полиномом*  $F(x)$ . Разрешенные кодовые слова циклического кода соответствуют полиномам, кратным порождающему полиному  $F(x)$ . При этом порождающий полином должен удовлетворять двум условиям: 1) быть неприводимым, т.е. не делиться ни на какой другой полином; 2) на него должен делиться без остатка двучлен вида  $x^n - 1$  (в данном случае имеется в виду обычная операция деления).

С введением операции деления полиномов появляется действие вычитания коэффициентов по модулю  $m$ . В случае, когда обычная разность коэффициентов положительна, никаких затруднений с получением ее значения по модулю  $m$  не возникает. Для получения же значения по модулю  $m$  от отрицательного числа к нему нужно прибавлять (один или несколько

раз) величину  $m$  до получения первого положительного значения или нуля, которые и принимаются за результат. Например,  $1\ominus_3 3 = 1$ ;  $1\ominus_8 3 = 6$ .

В частном случае двоичных кодов операция вычитания коэффициентов исключается, так как операции сложения и вычитания по модулю 2 эквивалентны:  $1\oplus 1 = 0$ ;  $1\ominus 1 = 0$ ;  $0\ominus 1 = 1$ . Это позволяет при использовании двоичных кодов при всех операциях над полиномами (включая и деление) пользоваться только суммированием коэффициентов по модулю 2 (двучлен  $x^n - 1$  при этом заменяется соответственно двучленом  $x^n + 1$ ).

Циклические коды являются блочными, равномерными и линейными. Линейность циклических кодов вытекает из того, что любая линейная композиция двух циклических кодов, делящихся на порождающий полином  $F(x)$ , будет обязательно делиться на тот же порождающий полином, т.е. будет принадлежать к числу разрешенных слов данного циклического кода. На разрешенный набор кодовых слов в этом случае накладывается дополнительное ограничение по сравнению с обычными линейными кодами: делимость на порождающий полином, и этот признак существенно упрощает аппаратную реализацию кодирующих и декодирующих устройств.

Базисная матрица циклического кода, являющегося линейным, может быть записана в виде:

$$L(x) = \begin{bmatrix} F(x) \\ xF(x) \\ \vdots \\ x^{k-1}F(x) \end{bmatrix}. \quad (9.30)$$

Матрицу (9.30), представленную в форме полиномов, можно переписать в форме матрицы коэффициентов (символов), записывая в одном столбце коэффициенты при одинаковой степени  $x$ . Наивысшая степень, отвечающая полиному в последней строке матрицы (9.30), равна  $k-1+r$ , где  $r$  — степень порождающего полинома  $F(x)$ . Поскольку строки этой матрицы, представляющие полиномы по модулю  $x^n - 1$ , не могут иметь степень выше  $n-1$ , то должно удовлетворяться условие  $k-1+r = n-1$  или  $k = n-r$ . Нетрудно убедиться, что все полиномы в (9.30) соответствуют разрешенному набору слов данного циклического кода [делятся на  $F(x)$ ] и являются линейно-независимыми, т.е. матрица (9.30), действительно, является базисной матрицей линейного кода  $(n, k)$ .

Проверочная матрица этого кода образуется аналогично, только вместо порождающего полинома  $F(x)$  в ней используется так называемый генераторный полином

$$g(x) = (x^n - 1)/F(x), \quad (9.31)$$

степень которого равна  $k$ . Она имеет следующий вид:

$$H(x) = \begin{bmatrix} g(x) \\ xg(x) \\ \vdots \\ x^{k-1}g(x) \end{bmatrix}. \quad (9.32)$$

В ортогональности полиномов, составляющих строки матриц  $L(x)$  и  $H(x)$ , легко убедиться, так как в силу (9.31)

$$x^i F(x) \otimes x^j g(x) = x^{i+j} \otimes (x^n - 1) = 0.$$

Можно показать, что если кодовое слово  $b = (\beta_0, \beta_1, \dots, \beta_{n-1})$  принадлежит к разрешенному набору данного циклического кода, то к нему же принадлежит и кодовое слово  $b^{(1)} = (\beta_{n-1}, \beta_0, \beta_1, \dots, \beta_{n-2})$ , полученное из первого циклической перестановкой символов. Легко убедиться, что кодовому слову  $b^{(1)}$  отвечает полином  $b^{(1)}(x)$ , представляющий произведение  $xb(x)$  по модулю  $x^n - 1$ . Поэтому, если  $b(x)$  делится на порождающий полином  $F(x)$ , то на него же делится и полином  $b^{(1)}(x)$ , и кодовое слово  $b^{(1)}$  принадлежит тому же циклическому коду, что и исходное слово  $b$ . Это свойство и определило название "циклический" для данного кода.

Формирование разрешенных слов циклического кода по заданному информационному слову, содержащему  $k$  символов, можно осуществить следующим образом. Образует полином  $G(x)$  степени  $k-1$ , соответствующий передаваемому  $k$ -разрядному информационному слову, и умножим его на  $x^r$ , где  $r$  — степень порождающего полинома  $F(x)$ ,  $k+r \leq n$ . Полученный полином степени  $k+r-1$  разделим на порождающий полином  $F(x)$ . Вычтем из полинома  $x^r G(x)$ , младший член которого имеет степень  $r$ , остаток  $q(x)$  от его деления на  $F(x)$  [порядок полинома  $q(x)$ , по крайней мере на единицу, ниже порядка полинома  $F(x)$ , т.е. не выше  $r-1$ ]. Получим полином  $Q(x) = x^r G(x) - q(x)$ , делящийся на  $F(x)$  без остатка, определяющий разрешенное кодовое слово. Младшие  $r$  разрядов этого кодового слова (проверочные) соответствуют коэффициентам остатка  $q(x)$ , а старшие разряды, начиная с  $(r+1)$ -го соответствуют информационному слову, смещенному на  $r$  разрядов в результате операции умножения полинома информационного слова  $G(x)$  на  $x^r$ .

Циклический код имеет систематическую форму, при этом все  $r$  проверочных символов идут подряд, занимая младшие разряды слова. Число его проверочных символов  $r$  определяется степенью порождающего полинома, а

общее число разрядов  $n$  — степень двучлена  $x^n - 1$ . Максимальное число информационных символов  $k = n - r$ .

В качестве примера определим разрешенное слово семизначного двоичного циклического кода, определяемое информационным словом 1011 и порождающим полиномом  $F(x) = x^3 + x^2 + 1$ , соответствующим двучлену  $x^3 + 1$ . Разделим полином  $x^7 G(x) = x^6 + x^4 + x^3$  на полином  $F(x)$  (напомним, что для двоичных кодов вычитание коэффициентов заменяется их суммированием по модулю 2):

$$\begin{array}{r} \oplus \begin{array}{r} x^6 + x^4 + x^3 \\ x^6 + x^5 + x^3 \\ \hline x^3 + x^4 + x^3 \\ \oplus x^5 + x^4 + x^2 \\ \hline x^2 \end{array} \quad \left| \begin{array}{r} x^3 + x^2 + 1 \\ x^3 + x^2 \\ \hline 1 \end{array} \right. \end{array}$$

Остаток  $q(x) = x^2$ . Следовательно,  $Q(x) = x^7 G(x) + q(x) = x^6 + x^4 + x^3 + x^2$ , а искомый циклический код, описываемый полиномом  $Q(x)$ , имеет вид 1011100 [отсутствующим в полиноме  $Q(x)$  степеням отвечают коэффициенты, равные нулю].

Для обнаружения ошибок в циклическом коде вместо операции деления на порождающий полином  $F(x)$  удобно использовать операцию умножения по модулю  $x^n - 1$  на генераторный полином  $g(x)$  (для двоичных кодов, когда используется только суммирование коэффициентов по модулю 2, умножение полиномов осуществляется соответственно по модулю  $x^n + 1$ ).

Полином  $\tilde{Q}(x)$ , соответствующий принятому коду, можно представить в виде суммы

$$\tilde{Q}(x) = Q(x) \oplus n(x), \quad (9.33)$$

где  $Q(x)$  — полином, соответствующий переданному циклическому коду;  $n(x)$  — полином, соответствующий "коду ошибки", имеющему единицы на месте искаженных символов и нули на остальных позициях.

Умножая  $\tilde{Q}(x)$  на генераторный полином  $g(x)$  по модулю  $x^n - 1$  (для двоичных кодов — по модулю  $x^n + 1$ ), получаем:

$$\tilde{Q}(x) \otimes_n g(x) = \frac{Q(x)}{F(x)} \otimes_n (x^n - 1) + n(x) \otimes_n g(x) = n(x) \otimes_n g(x), \quad (9.34)$$

так как  $g(x) = (x^n - 1)/F(x)$ , а  $Q(x)$  делится без остатка на  $F(x)$  по определению циклического кода. Произведение  $n(x) \otimes_n g(x)$  не равно нулю, если только "код ошибки" не совпадает ни с одним из разрешенных слов цик-

лического кода данного типа. Таким образом, признаком наличия ошибки является отличие от нуля произведения  $Q(x) \otimes_n g(x)$ ; при этом обнаруживается ошибка любой кратности, если только она не совпадает с одним из разрешенных слов данного циклического кода.

Умножение по модулю  $x^n + 1$  для двоичных кодов легко реализуется с использованием кольцевого регистра из  $n$  ячеек с сумматором по модулю 2 на входе (рис. 9.1). Пусть на вход кольцевого регистра последовательно поступают, начиная с младших разрядов, коэффициенты обычного произведения полиномов  $\tilde{Q}(x)g(x)$ . Представим это произведение в виде

$$Q(x)g(x) = p(x)(x^n + 1) + n(x) \otimes_n g(x), \quad (9.35)$$

разбивая его на часть, кратную  $x^n + 1$ , и остаток  $n(x) \otimes_n g(x)$ .

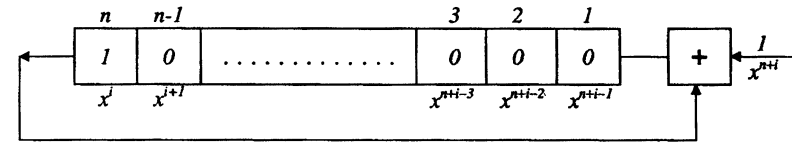


Рис. 9.1

Первое произведение в правой части (9.35) разбивается на пары членов вида  $x^{n+i} + x^i$ , которые взаимно "гасятся" при прохождении через кольцевой регистр. Действительно, в тот момент, когда на входной сумматор (рис. 9.1) поступает коэффициент 1 при члене  $x^{n+i}$ , по цепи обратной связи из  $n$ -й ячейки регистра в него поступает коэффициент 1 при  $x^i$ . Суммируясь по модулю 2, они гасятся. В результате в момент поступления в регистр последнего (старшего) члена полинома произведения (9.35) в кольцевом регистре будут зафиксированы коэффициенты полинома, равного

$$n(x) \otimes_n g(x).$$

Отсутствие единиц в кольцевом регистре после окончания этой процедуры говорит о том, что ошибки в принятом кодовом слове либо отсутствуют, либо полином ошибки кратен порождающему полиному  $F(x)$ ; наличие же единиц служит критерием искажений данного кодового слова. Таким образом, рассмотренная весьма простая схема позволяет обнаружить наличие ошибок, не совпадающих с разрешенными кодовыми словами.

Исправление ошибок, т.е. определение полинома ошибки  $n(x)$  возможно лишь по отношению к ошибкам заданной структуры, ибо в общем случае знание произведения  $n(x) \otimes_n g(x)$  не позволяет однозначно восстановить значение соответствующего обычного произведения  $n(x)g(x)$  и определить  $n(x)$ . Структуру ошибки можно задавать числом искаженных символов и интервалами между ними. Ошибка заданной структуры полностью определяется положением младшего из искаженных разрядов кодового слова. Ее полином можно представить в виде

$$n(x) = x^{i-1} n_0(x), \quad (9.36)$$

где  $\mu$  — номер младшего из искаженных разрядов;  $n_0(x)$  — ошибка той же структуры, начинающаяся с первого разряда.

Остаток в кольцевом регистре, отвечающий ошибке (9.36), равен

$$x^{\mu-1}[n_0(x) \otimes g(x)], \quad (9.37)$$

т.е. отличается от  $n_0(x) \otimes g(x)$  сдвигом вправо (в сторону старших разрядов) на  $\mu - 1$  разрядов регистра. Положение ошибки определяется, таким образом, числом разрядов, на которое нужно продвинуть остаток в кольцевом регистре, чтобы он совпал с "эталонным" полиномом  $n_0(x) \otimes g(x)$ , хранящимся в памяти. В частности, для одиночных ошибок "эталонном" для сравнения служит генераторный полином  $g(x)$ .

К циклическим кодам принадлежит и  $m$ -последовательность, рассматривавшаяся в § 4.6. Введем вместо символов  $+1$  и  $-1$ , использовавшихся в § 4.6 при изучении корреляционных свойств ФМПС, символы  $1$  и  $0$ . При переходе к новым символам (обозначим их  $d_i$ ) закон (4.61) формирования  $m$ -последовательности трансформируется к виду

$$d_i = d_{i-r} \oplus d_{i-l} \oplus \dots \oplus d_{i-k}, \quad (9.38)$$

где по-прежнему полагается, что среди символов  $d_{i-1}, d_{i-2}, \dots, d_{i-r}$  хотя бы один равен  $1$ . Легко убедиться в эквивалентности закона формирования символов (9.42) использовавшемуся в § 4.6 закону  $d_i = -d_{i-r}d_{i-l} \dots d_{i-k}$  при замене символа  $-1$  на  $0$ .

Используя представление кодов в виде полиномов, закон (9.38) можно записать в виде:

$$d_i x^{i-1} = d_{i-r} x^{i-r-1} x^r + d_{i-l} x^{i-l-1} x^l + \dots + d_{i-k} x^{i-k-1} x^k. \quad (9.39)$$

Из (9.39) следует, что формирование полинома, отвечающего  $m$ -последовательности, можно рассматривать как процесс "наращивания" старших членов этого полинома путем умножения предшествующих его членов на полином

$$F(x) = x^r + x^l + \dots + x^k. \quad (9.40)$$

Покажем, что символы, составляющие период  $m$ -последовательности, формируемой по закону (9.38), могут рассматриваться как  $N$ -разрядный двоичный циклический код с порождающим полиномом (9.40). Рассмотрим полином степени  $N + r - 1$ :

$$Q_{N+r}(x) = \sum_{i=r+1}^{N+r} d_i x^{i-1}, \quad (9.41)$$

коэффициенты которого составляют период  $m$ -последовательности.

Учитывая, что  $d_{N+j} = d_j$  (в силу периодичности  $m$ -последовательности) и что

$$d_{N+j} x^{N+j-1} \pmod{(x^N + 1)} = d_j x^{j-1} = d_j x^{j-1}, \quad (9.42)$$

получаем:

$$\begin{aligned} Q_{N+r}(x) \pmod{(x^N + 1)} &= d^{r+1} x^r + \dots + d_N x^{N-1} + \\ &+ d_1 + d_2 x + \dots + d_{r-1} x^{r-1} = Q_N(x), \end{aligned} \quad (9.43)$$

где  $Q_N(x)$  — полином, отображающий  $N$ -разрядную кодую последовательность

$$b = (d'_1, d'_2, \dots, d'_r, d'^{r+1}, \dots, d'_N). \quad (9.44)$$

Используя (9.39) и (9.42), полином (9.41) представим в виде:

$$\begin{aligned} Q_{N+r} &= x^r \sum_{i=r+1}^{N+r} d_{i-r} x^{i-r-1} + x^l \sum_{i=r+1}^{N+r} d_{i-l} x^{i-l-1} + \dots + \\ &+ x^k \sum_{i=r+1}^{N+r} d_{i-k} x^{i-k-1} = x^r \sum_{i=1}^N d_i x^{i-1} + x^l \sum_{i=1}^N d_i x^{i-1} + \dots + \\ &+ x^k \sum_{i=1}^N d_i x^{i-1} = F(x) \sum_{i=1}^N d_i x^{i-1}. \end{aligned} \quad (9.45)$$

Таким образом, полином  $Q_{N+r}(x)$ , а значит, и полином  $Q_N(x) = Q_{N+r}(x) \pmod{x^N + 1}$  кратен порождающему полиному (9.40) и последовательность (9.44) представляет циклический код. Любой другой участок  $m$ -последовательности длительностью в один период  $d'_1 \dots d'_N d'_{N+1} \dots d'_{N+l} = d'_1 \dots d'_N d'_1 \dots d'^{l-1}$ , который может быть получен циклической перестановкой последовательности (9.44), принадлежит тому же циклическому коду.

Теперь можно уточнить критерий выбора значений индексов  $r, l, \dots, k$  в законах (4.61) и (9.38) формирования символов  $m$ -последовательности: эти индексы определяются порождающим полиномом  $F(x)$ . При этом, для получения последовательности максимальной длительности порождающий полином должен обладать дополнительным свойством "примитивности": двучлен  $x^N - 1$  должен делиться на порождающий полином степени  $r$  лишь при  $n = N = 2^r - 1$  и не должен делиться ни при каком значении  $n < N$ .

Циклические коды типа  $m$ -последовательности, формируемые по закону (9.38), составляют лишь небольшое подмножество всех кодовых слов, принадлежащих данному циклическому коду: из общего числа  $2^{N-r}$  разрешенных кодовых слов лишь  $N$  кодовых слов представляют  $m$ -последовательности. Из каждых  $2^{N-r}/2^r = 2^{N-2r}$  кодовых слов с данным сочетанием проверочных символов лишь одно представляет  $m$ -последовательность. Из свойств  $m$ -последовательности, рассмотренных в § 4.6 [см. (4.62)], следует, что разрешенные кодовые слова, принадлежащие  $m$ -последовательности, отличаются друг от друга ровно на  $(N+1)/2 = 2^{r-1}$  символов, что отвечает максимально возможному кодовому расстоянию для разрешенного набора, содержащего  $N$   $N$ -разрядных слов.

## 9.4. Критерии эффективности блочных корректирующих кодов

В качестве критерия эффективности корректирующих кодов можно использовать обеспечиваемую ими достоверность передачи информации при фиксированных скорости передачи информации, средней мощности сигнала и помехах.

Для представления этого критерия необходимо прежде всего ввести меру достоверности передачи информации, обеспечивающую возможность

сравнения по этой характеристике различных кодов. Для блочных корректирующих кодов в качестве такой меры удобно использовать предложенную Финком эквивалентную вероятность ошибки  $P_3$ , характеризующуюся значением вероятности ошибочного приема элементарного символа примитивного кода  $(k, k)$ , при котором последний обеспечивает ту же достоверность передачи информации, что и рассматриваемый блочный корректирующий код  $(n, k)$ . Поскольку для примитивного кода вероятность ошибочного приема символа при случайных независимых искажениях отдельных символов однозначно характеризует достоверность передачи информации, то величина  $P_3$  является однозначной характеристикой достоверности корректирующего кода.

Если оцениваемый код  $(n, k)$  в рассматриваемом канале характеризовать вероятностью ошибочного приема  $P_k$ , то вероятность  $P_3$  ошибочного приема символа, при которой обеспечивается та же достоверность передачи информации для примитивного кода  $(k, k)$ , определяются из равенства  $(1 - P_3)^k = 1 - P_k$  или

$$P_3 = 1 - (1 - P_k)^{1/k}. \quad (9.46)$$

При  $P_k \ll 1$

$$p_3 \approx P_k/k. \quad (9.47)$$

Если корректирующий код  $(n, k)$  имеет кодовое расстояние  $d_{\min}$ , т.е. способен достоверно исправить  $t = (d_{\min} - 1)/2$  ошибок, то при случайных независимых искажениях его символов, характеризующихся вероятностью  $P_e$ , получим:

$$P_k = \sum_{i=t+1}^N C_n^i P_e^i (1 - P_e)^{n-i}, \quad (9.48)$$

$$P_3 = 1 - (1 - P_k)^{1/k} = 1 - \left[ \sum_{i=t+1}^n C_n^i P_e^i (1 - P_e)^{n-i} \right]^{1/k}. \quad (9.49)$$

В случае, когда

$$P_3 \ll 1 \text{ и } P_k = \sum_{i=t+1}^n C_n^i P_e^i (1 - P_e)^{n-i} \ll 1,$$

сохраняя в (9.49) лишь первый член суммы, имеющий наименьший  $(t+1)$ -й порядок относительно малой величины  $P_e$ , получаем:

$$P_3 \approx \frac{P_k}{k} \approx \frac{C_n^{t+1} P_e^{t+1} (1 - P_e)^{n-t-1}}{k} \approx \frac{C_n^{t+1} P_e^{t+1}}{k}. \quad (9.50)$$

Для сравнения двух кодов  $(n_i, k_i)$  и  $(n_j, k_j)$ , имеющих разрешенные наборы кодовых слов  $B_{P_i}^{(i)}$  и  $B_{P_j}^{(j)}$ , можно воспользоваться отношением

$$\beta_{ji} = \frac{P_{3i}}{P_{3j}} \approx \frac{k_j C_{n_i}^{t_i+1} P_{e_i}^{t_i+1}}{k_i C_{n_j}^{t_j+1} P_{e_j}^{t_j+1}}, \quad (9.51)$$

где  $\beta_{ji}$  — коэффициент, характеризующий выигрыш  $j$ -го кода по отношению к  $i$ -му;  $t_i$  и  $t_j$  — число достоверно исправляемых ошибок, соответствующие кодовым расстояниям сравниваемых кодов  $d_{\min i}$  и  $d_{\min j}$ ;  $P_{e_i}$  и  $P_{e_j}$  — вероятности искажений элементарных символов сравниваемых кодов при одинаковых значениях скорости передачи информации, средней мощности канала и характеристиках помех.

Рассмотрим частный случай, когда помеха — аддитивный белый шум со спектральной плотностью  $N_0$ . Скорость передачи информации зафиксируем числом  $m$  информационных символов, передаваемых в 1 с (исходный информационный код считаем безызбыточным). Энергию обоих символов двоичного кода принимаем одинаковой. Тогда, используя формулу (3.38), получаем:

$$P_{e_i} = F\left(-\sqrt{\frac{\lambda E_i}{N_0}}\right) = F\left(-\sqrt{\frac{\lambda \mathcal{P}_c}{N_0 m n_i / k_i}}\right), \quad (9.52)$$

где  $\mathcal{P}_c$  — фиксируемая средняя мощность канала;  $m n_i / k_i$  — общее число символов, передаваемых за 1 с при применении кода  $(n_i, k_i)$ , отвечающее условию передачи  $m$  информационных символов в 1 с;  $\lambda$  — коэффициент различимости символов (3.34).

В качестве характеристик эффективности корректирующего кода удобно использовать значение коэффициента выигрыша  $\beta$  по отношению к примитивному коду. Для примитивного кода  $t_i = t_{\text{пр}} = 0$ ,  $P_{3i} = P_{\text{епр}}$ . Поэтому выигрыш  $\beta$  по эквивалентной вероятности ошибки корректирующего кода  $(n, k)$ , достоверно исправляющего  $t$  ошибок, по отношению к примитивному  $k$ -символьному коду в соответствии с (9.51) равен

$$\beta = P_{\text{епр}} / P_{3\text{кор}} \approx k P_{\text{епр}} / C_n^{t+1} P_{\text{екор}}^{t+1}, \quad (9.53)$$

где  $P_{\text{екор}}$  — вероятность ошибочного приема символа корректирующего кода;  $P_{\text{епр}}$  — вероятность ошибочного приема символа соответствующего примитивного кода при тех же помехах и тех же значениях скорости передачи информации и средней мощности канала.

В частности, для помехи типа аддитивного белого шума

$$\beta \equiv \frac{F\left(-\sqrt{\frac{\lambda E_c n}{N_0 k}}\right)}{\frac{C_n^{t+1}}{k} \left[F\left(-\sqrt{\frac{\lambda E_c}{N_0}}\right)\right]^{t+1}}, \quad (9.54)$$

где  $E_c$  — энергия символа корректирующего кода ( $n, k$ );  $E_{cn}/k$  энергия символа примитивного кода при тех же значениях скорости передачи информации и средней мощности сигнала;  $\lambda$  — коэффициент различимости символов (см. (3.34)).

В качестве примера на рис. 9.2 приведен график зависимости коэффициента выигрыша  $\beta$  от отношения  $E_c/N_0$  для 15-разрядного кода Хэмминга, исправляющего одиночные ошибки ( $n = 15, k = 11, t = 1$ ). График построен для  $\lambda = 2$  в области достаточно больших значений  $E_c/N_0$ , когда применима приближенная формула (9.54). Эффективность корректирующих кодов растет с увеличением достоверности приема отдельных символов. При низкой же достоверности приема символов (малых значениях  $E_c/N_0$ ) повышение корректирующей способности кода может приводить даже к отрицательным результатам, так как при этом преобладающим может оказаться увеличение вероятности ошибочного приема символов из-за снижения их энергии при увеличении избыточности.

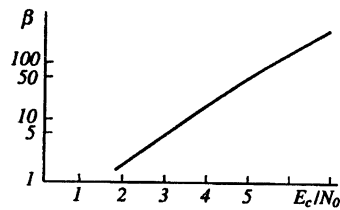


Рис. 9.2

## Глава 10

### СВЕРТОЧНЫЕ КОРРЕКТИРУЮЩИЕ КОДЫ

#### 10.1. Особенности кодирования и декодирования сверточных кодов

Ограничимся рассмотрением двоичных сверточных кодов. Для них рекуррентная формула (8.4) для определения символов кодовой последовательности принимает вид

$$\beta_{ij} = \left\{ \sum_{v=q}^{s-1} C_{vj} b_{i+v} \right\} (\text{mod } 2); \quad j = 0, 1, \dots, l-1, \quad c_{vj} = 0, 1, \quad (10.1)$$

где  $b_i$  — символы входной информационной последовательности.

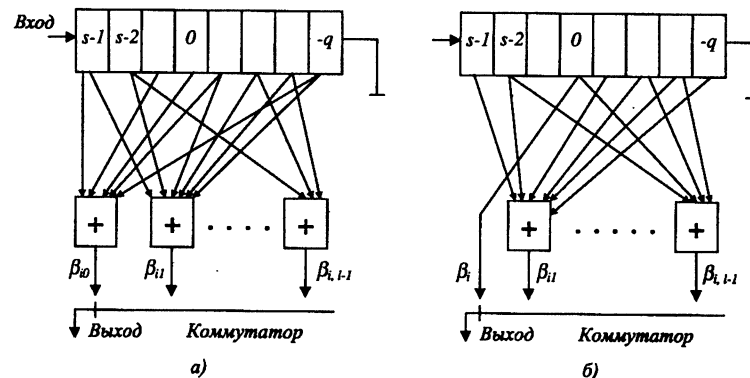


Рис. 10.1

Функциональная схема кодирующего устройства, обеспечивающего формирование символов по закону (10.1), представлена на рис. 10.1, а. Символы входной информационной последовательности поступают на регистр сдвига, содержащий  $k = s + q$  ячеек [в соответствии с обозначениями в (10.1) ячейкам присвоены номера от  $-q$  до  $s-1$ ]. Символы  $\beta_{ij}$  формируются с помощью сумматоров по модулю 2, на которые с ячеек регистра сдвига, отвечающих ненулевым значениям  $c_{vj}$  в (10.1), поступают информационные символы. В тот такт работы регистра сдвига, когда в нулевой его ячейке находится символ  $b_i$  (соответственно во входной ячейке символ

$b_{i+q-1}$ , а в выходной —  $b_{i-q}$ , на сумматорах формируются символы  $\beta_{i0}$ ,  $\beta_{i1}$ , ...,  $\beta_{i, l-1}$ . Коммутатор в  $i$ -м такте работы регистра сдвига, в течение которого остаются неизменными заполняющие его символы  $b_i$ , поочередно подает символы  $\beta_{ij}$  в канал. В следующем такте все символы в регистре сдвинуты вправо на одну ячейку; в нулевой ячейке будет символ  $b_{i+1}$ , во входную ячейку поступит символ  $b_{i+q}$ , а в выходную —  $b_{i-q+1}$ . Соответственно на сумматорах формируются символы кодовой последовательности  $\beta_{i+1,0}$ ,  $\beta_{i+1,1}$ , ...,  $\beta_{i+1, l-1}$  и т.д. Таким образом, с поступлением на вход кодирующего устройства каждого нового символа информационной последовательности в канал поступают  $l$  символов кодовой последовательности.

На рис. 10.1,б представлена модификация кодирующего устройства для случая систематического сверточного кода. Отличие в этом случае состоит в том, что на один из входов коммутатора (например, на первый) непосредственно из нулевой ячейки регистра поступает информационный символ  $b_i$ , т.е.  $\beta_{i0} = b_i$ . Основное отличие от кодирования блочных систематических кодов здесь состоит в непрерывном обновлении группы информационных символов, участвующих в формировании очередных  $l$  символов кодовой последовательности (при блочном кодировании формирование всех символов данного кодового слова происходит с использованием одного кодового слова происходит с использованием одного и того же блока информационных символов, который целиком обновляется при переходе к кодированию следующего слова). Максимальное число информационных символов, участвующих в формировании каждого выходного символа сверточного кода, определяемое числом ячеек регистра сдвига  $K$ , носит название *длины кодовых ограничений*. Эта характеристика близка по смыслу к длине блока информационных символов для блочных кодов. Иногда длину кодовых ограничений определяют в числе канальных (выходных) символов. В этом случае она выражается величиной  $N = lK$ . Каждый символ входной информационной последовательности может участвовать в формировании до  $lK - 1$  символов выходной последовательности сверточного кода. Нетрудно видеть, что усложнение кодирующего устройства с увеличением длины кодовых ограничений характеризуется линейным законом. Поэтому кодирующие устройств не накладывают существенных ограничений на реализуемость сверточных кодов.

По способу декодирования сверточные коды можно разделить на два класса: коды с синдромной коррекцией и коды с использованием последовательного декодирования.

В качестве кодов с синдромной коррекцией используются только систематические сверточные коды. Способ коррекции основывается на образовании из принятой кодовой последовательности контрольных сумм вида

$$S_{ij} = \beta'_{ij} \oplus \sum_{v=q}^{s-1} c_{vj} \beta'_{i+v}, \quad j = 1, 2, \dots, l-1, \quad (10.2)$$

в каждую из которых, как и в случае блочных линейных кодов, входит один из проверочных символов  $\beta'_{ij}$ . Принцип коррекции такой же, как и для блочных линейных кодов: законы формирования контрольных символов и контрольных сумм выбираются так, чтобы синдром ошибки, составленный из значений контрольных сумм, позволял определить искаженные символы. При отсутствии искажений все контрольные суммы равны нулю, а синдром ошибки представляет нуль-вектор.

Сверточные коды с использованием последовательного декодирования характеризуются отсутствием каких-либо ограничений как на структуру связей между символами информационной и кодовой последовательностей, так и на структуру исправляемых ошибок (в пределах корректирующей способности кода). Это создает принципиальную возможность, используя несистематические коды, получать выигрыш в эффективности кодов более чем в  $1/(1 - p_n)$  раз. При этом выигрыш в  $1/(1 - p_n)$  раз достигается благодаря тому, что в несистематическом коде все символы кодовой последовательности (а не только избыточные) участвуют в коррекции ошибок, а дополнительный выигрыш — благодаря отсутствию ограничения на выбор их связей с информационными символами. Однако практическая реализация этого выигрыша возможна лишь при создании алгоритмов декодирования несистематических сверточных кодов произвольной структуры, вычислительная сложность которых не выходила бы за возможности существующих цифровых ЭВМ.

В самом общем виде схему декодирования сверточных кодов произвольной структуры можно было бы представить следующим образом. Для всех возможных реализаций входной информационной последовательности по формулам (10.1) определяются соответствующие разрешенные кодовые последовательности. Декодирование заключается в отборе наиболее правдоподобных из них по отношению к принятой кодовой последовательности (например, самой близкой к принятой последовательности по метрике Хэмминга). Однако такой "лобовой" метод очень громоздок и практически непригоден (для информационной последовательности, содержащей  $n$  символов, потребовалось бы рассмотрение  $2^n$  разрешенных кодовых последовательностей). Поэтому для создания реализуемых алгоритмов декодирования несистематических сверточных кодов произвольной структуры необходимо найти такой подход к решению задачи, который позволил бы резко сократить объем перебора разрешенных кодовых последовательностей при поиске наиболее правдоподобной из них, так чтобы сложность декодирующего устройства возрастала бы как малая степень длины кодовых ограничений.



Для отображения вариантов перебора удобно пользоваться "деревом информационных последовательностей". На рис. 10.2 представлено дерево, отображающее все возможные реализации входной информационной последовательности. От его основания идут две ветви, соответствующие двум возможным значениям (0 или 1) первого символа информационной последовательности. Каждая из них дает две ветви, соответствующие возможным значениям второго символа, и т.д. Варианты перебора определяются возможными "траекториями" перемещения по ветвям дерева информационных последовательностей, начиная от его основания.

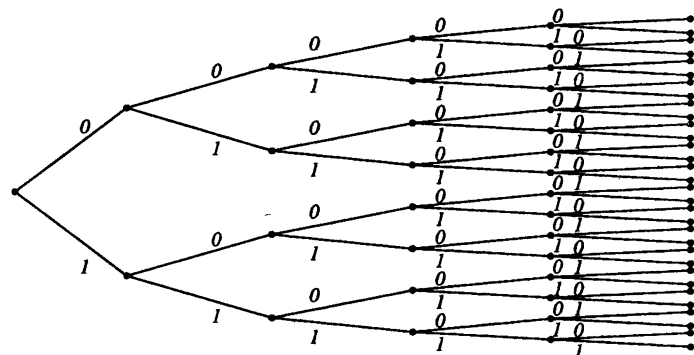


Рис. 10.2

Возможность резкого сокращения объема перебора кодовых последовательностей при декодировании несистематических сверточных кодов определяется следующими двумя способами:

1. Корреляционные связи между символами кодовой последовательности ограничены областью, определяемой длиной кодовых ограничений  $K$ . Поэтому, если вместо распознавания всего принятого сообщения, осуществлять последовательное распознавание отдельных его участков, имеющих длительность, соизмеримую с длиной кодовых ограничений, то для первых символов выделенного участка последовательности будут учтены все их связи с остальными символами и достоверность их распознавания практически не ухудшится. Это позволяет реализовать процедуру декодирования в виде последовательного распознавания этих участков с принятием на каждом таком участке окончательного решения лишь в отношении первого символа и смещением каждый раз распознаваемого участка на один символ. Поскольку количество разрешенных комбинаций определяется показательной зависимостью от длительности рассматри-

ваемого участка информационной последовательности, общий объем подлежащих перебору последовательностей при этом резко сокращается.

2. Подавляющее большинство из числа возможных комбинаций информационных символов являются маловероятными и при сравнении отвечающих им кодовых последовательностей с принятой последовательностью могут быть отброшены после формирования первых же символов. Это также приводит к резкому сокращению объема перебора (для большинства вариантов рассмотрение прекращается уже в самом начале оцениваемого участка траектории перемещения по дереву информационных последовательностей).

Эти идеи и лежат в основе метода последовательного декодирования, впервые предложенного Возенкрафтом [29]. Появление метода последовательного декодирования привело к резкому увеличению интереса к сверточным кодам в теории и практике кодирования: предлагаются новые более экономные алгоритмы последовательного декодирования, исследуется их эффективность. Общие идеи некоторых из этих алгоритмов рассматриваются в § 10.3.

## 10.2. Сверточные коды с синдромной коррекцией

Существенной особенностью сверточных кодов с синдромной коррекцией является большая по сравнению с блочными кодами свобода выбора информационных символов, участвующих в формировании проверочных символов. Если для блочных кодов проверочные символы и контролируемые ими информационные символы должны быть сосредоточены в рамках кодового слова, то в сверточных кодах имеется возможность их "рассредоточить" (по этому признаку блочные коды относятся к *сосредоточенным*, а сверточные к *рассредоточенным* кодам). В случае действия некоррелированных помех, когда искажения отдельных символов могут рассматриваться как случайные независимые события, различие кодов по этому признаку не является существенным. Однако в случае воздействия на канал коррелированной помехи, приводящей к появлению пачки ошибок, затрагивающей целую цепочку символов (обычно это сильные кратковременные помехи типа разряда молнии, кратковременных замираний сигнала и т.п.), рассредоточенные коды обладают существенными преимуществами. Если обеспечить достаточный разнос между информационными и отвечающими им проверочными символами, так чтобы они не подвергались одновременному воздействию пачки ошибок, то становится возможным исправить все ошибки в пределах этой пачки, независимо от их кратности. Поэтому в тех случаях, когда по условиям связи приходится

считаться с возможностью действия пачек ошибок, сверточные коды с синдромной коррекцией весьма эффективны.

В качестве примера рассмотрим простейший двоичный сверточный код Финка—Хагельбергера, широко применяемый в практике связи. Это систематический сверточный код с избыточностью 0,5 следующей структуры

$$\dots b_{i-h}\beta_{i-h} \dots b_{i-1}\beta_{i-1}b_i\beta_i \dots b_{i+h+1}\beta_{i+h+1} \dots \quad (10.3)$$

Формула для вычисления проверочных символов этого кода имеет вид:

$$\beta_i = b_{i-h} \oplus b_{i+h+1} \quad (10.4)$$

и является частным случаем формулы (10.1), когда коэффициенты  $c_{vj}$  отличны от нуля только при  $v = -h$  и  $v = h+1$ . Составляющие синдрома ошибки для этого кода в соответствии с (10.2) равны

$$S_i = \beta'_i \oplus b'_{i-h} \oplus b'_{i+h+1}. \quad (10.5)$$

Из (10.3) и (10.4) видно, что проверочные символы  $\beta_i$  получаются суммированием симметрично расположенных относительно них информационных символов кодовой последовательности, отстоящих на  $2h+1$  позиций в каждую сторону.

$$\dots b'_{i-2h-1} \dots \beta'_{i-h-1} \dots b'_i \dots b'_{i+h} \dots b'_{i+2h+1} \dots$$

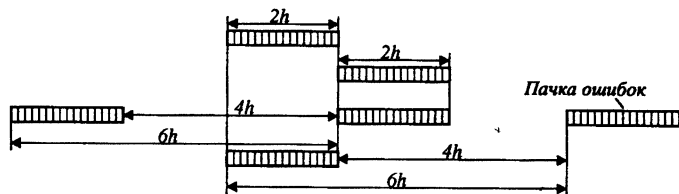


Рис. 10.3

Выберем параметр  $h$  так, чтобы выполнялись следующие условия: если информационный символ  $b_i$  подвергся воздействию пачки ошибок, то ни данная, ни соседние пачки ошибок не должны захватывать ни проверочные символы  $\beta_{i+h}$  и  $\beta_{i-h-1}$ , в формировании которых участвует символ  $b_i$  ни информационные символы  $b_{i+2h+1}$  и  $b_{i-2h-1}$ , участвующие в формировании тех же проверочных символов. Как видно из рис. 10.3, для того чтобы одна и та же пачка ошибок одновременно с символом  $b'_i$  не могла захватить и символ  $\beta'_{i+h}$  или  $\beta'_{i-h-1}$ , ее максимальная длина  $l_{\max}$  должна удовлетворять условию:

$$l_{\max} \leq 2h. \quad (10.6)$$

Для того чтобы конец предыдущей и начало следующей пачки ошибок не могли захватывать символы  $b'_{i-2h-1}$  и  $b'_i$  или  $b'_i$  и  $b'_{i+2h+1}$  минимальный интервал между началами соседних пачек  $L_{\min}$  должен удовлетворять условию

$$L_{\min} \geq 6h + 1. \quad (10.7)$$

При выполнении условий (10.6) и (10.7) искажение информационного символа  $b'_i$  будет сопровождаться одновременным обращением в единицу пары составляющих синдрома ошибок:

$$S_{i-h-1} = \beta'_{i-h-1} \oplus b'_{i-2h-1} \oplus b'_i, \quad (10.8)$$

$$S_{i+h} = \beta'_{i+h} \oplus b'_i \oplus b'_{i+2h+1}, \quad (10.9)$$

индексы которых отличаются на  $2h+1$  [остальные символы в правых частях (10.8) и (10.9) не будут подвержены действию пачек ошибок]. Нетрудно убедиться, что при соблюдении условий (10.6) и (10.7) искажение проверочного символа вызовет обращение в единицу лишь составляющей синдрома с тем же индексом, составляющие же синдрома с индексом, отличающимися в ту или иную сторону на  $2h+1$ , сохраняют нулевое значение.

Таким образом, правила коррекции кода Финка—Хагельбергера сводятся к следующему: если отлична от нуля составляющая синдрома  $S_j$ , а составляющая  $S_{j-2h-1}$  равна нулю, то исправление ранее принятых информационных символов не производится; если наряду с составляющей  $S_j$  не равна нулю и составляющая  $S_{j-2h-1}$ , то исправляется информационный символ  $b'_{j-h}$ .

Код Финка—Хагельбергера позволяет исправить пачки ошибок, если только минимальный интервал между ними более чем в три раза превышает их максимальную длину [в противном случае невозможно выбрать параметр  $h$  кода, удовлетворяющий одновременно условиям (10.6) и (10.7)].

### 10.3. Сверточные коды с использованием последовательного декодирования

Если при синдромной коррекции область эффективного применения сверточных кодов ограничивалась лишь каналами, подверженными действию пачек ошибок, то появление алгоритмов последовательного декодирования существенно расширило область применения сверточных кодов. В настоящее время предложены алгоритмы последовательного декодирования, вычислительная сложность которых не выходит за рамки существующих технических возможно-

стей даже для сверточных кодов с достаточно большой длиной кодового ограничения. Эффективность таких кодов существенно выше эффективности блочных кодов и довольно близко подходит к шенноновскому пределу.

Сущность всех алгоритмов последовательного декодирования состоит в целенаправленном поиске "траектории" перемещения по дереву информационных последовательностей, изображенному на рис. 10.2, которая отвечала бы наиболее правдоподобной реализации информационной последовательности. Под целенаправленностью здесь понимается такой способ поиска этой траектории, который требует перебора лишь небольшой части ветвей этого дерева. Мерой правдоподобия траектории при посимвольном приеме может служить расстояние по Хэммингу между отвечающей этой траектории реализацией кодовой последовательности и фактически принятой последовательностью символов. По мере продвижения по рассматриваемой траектории и наращивания числа символов в оцениваемой кодовой последовательности расстояние между ней и соответствующим участком принятой последовательности при неверном выборе траектории будет быстро увеличиваться. Это позволяет прекращать движение по данной траектории, как только расстояние достигнет установленного порогового значения  $d_1$  (пороговое значение  $d_1$  по мере продвижения по траектории увеличивается).

Существуют различные подходы к сужению области перебора реализаций при последовательном декодировании. Мы ограничимся рассмотрением основных идей двух характерных с этой точки зрения алгоритмов: Возенкрафта [29] и Зигангирова—Джемника [28].

Ключевой позицией принятой Возенкрафтом стратегии последовательного декодирования является замена задачи поиска наиболее вероятной из всех разрешенных кодовых последовательностей задачей исключения всех маловероятных кодовых последовательностей, которые могут отвечать принятой последовательности лишь при большом числе искажений [29]. Каждый шаг предложенной Возенкрафтом процедуры последовательного декодирования заключается в переборе отрезков траекторий, которые могут служить продолжением выбранного на предшествующих шагах начального ее участка и закреплении первой ветви отобранного продолжения. Устанавливаются максимальная длина  $n$  отрезков траектории, рассматриваемых на каждом шаге декодирования (глубина перебора), и порог  $d_1$ , определяющий

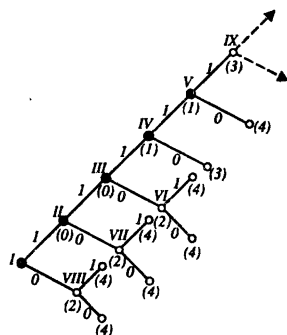


Рис. 10.4

допустимое расстояние между соответствующими участками рассматриваемой и принятой кодовых последовательностей. Отдельные варианты отсеиваются, как только это расстояние превышает  $d_1$ . Перебор прекращается при получении первого же варианта отрезка траектории длиной  $n$ , для которого это расстояние не превышает пороговый уровень  $d_1$ , после чего первая его ветвь (символ) включается в декодированный начальный участок траектории и вся процедура выбора продолжения траектории повторяется от новой узловой точки. Если на каком-либо шаге декодирования при пороге  $d_1$  оказываются отсеянными все варианты продолжений траектории заданной длины  $n$ , процедура на данном шаге повторяется при

новом значении порога  $d_2 > d_1$ . Вычислительная сложность данного алгоритма последовательного декодирования определяется выбором глубины перебора  $n$ , значений пороговых уровней  $d_1, d_2, \dots$  и числом градаций пороговых уровней (обычно ограничиваются двумя градациями).

В алгоритме Зигангирова—Джемника [28] все внимание сосредоточено именно на определении оптимальной траектории. Рассмотрим основную идею алгоритма Зигангирова—Джемника. Для простоты опустим рассмотрение процедуры определения расстояния по Хэммингу между кодовой последовательностью, отвечающей рассматриваемой информационной последовательности, и фактически принятой последовательностью, а каждому варианту траектории перемещения по дереву информационных последовательностей будем условно приписывать некоторое значение  $d_i$ . В каждой узловой точке, начиная от основания дерева информационных последовательностей из двух возможных вариантов продолжения траектории выбирается тот, который обеспечивает меньшее удаление  $d_i$  соответствующей кодовой последовательности (включая и ранее определенный начальный ее участок) от принятой последовательности. На каждом шаге процедуры значение  $d_i$  и номер  $i$  наилучшего варианта начального участка траектории заносятся в верхнюю строку таблицы, хранящейся в памяти; остальные строки этой таблицы заполняются в порядке возрастания  $d_i$  начальными участками траекторий, соответствующими всем рассмотренным вариантам. Для нумерации участков траекторий удобно использовать соответствующие информационные последовательности (эти же номера относятся и к конечным узловым точкам участков). Если после очередного шага процедуры значение  $d_i$  для лучшего из двух продолжений траектории превысит хотя бы одно из хранящихся в памяти значений  $d_i$  для рассмотренных ранее вариантов, дальнейший поиск оптимального продолжения траектории переносится в  $j$ -й узел (при этом производится перестановка строк в таблице по возрастающим значениям  $d_i$ ). При равенстве значений  $d_i$  для двух вариантов траектории предпочтение отдается варианту с большей длиной. После  $t$  шагов процедуры (каждый шаг заключается в рассмотрении двух продолжений траектории из данного узла на одну ветвь) таблица будет содержать  $t + 1$  строку, максимальное же продвижение по траектории может быть существенно меньшим, чем на  $t$  ветвей, из-за возможных возвратов по дереву информационных последовательностей в процессе поиска оптимальных продолжений.

Таблица 10.1

| Номер строки | Номер шага |       |        |         |          |          |          |          |
|--------------|------------|-------|--------|---------|----------|----------|----------|----------|
|              | I          | II    | III    | IV      | V        | VI       | VII      | VIII     |
| 1            | 1(0)       | 11(0) | 111(1) | 1111(1) | 110(2)   | 10(2)    | 0(2)     | 11111(3) |
| 2            | 0(2)       | 10(2) | 110(2) | 110(2)  | 10(2)    | 0(2)     | 11111(3) | 1110(3)  |
| 3            |            | 0(2)  | 10(2)  | 10(2)   | 0(2)     | 11111(3) | 1110(3)  | 11110(4) |
| 4            |            |       | 0(2)   | 0(2)    | 11111(3) | 1110(3)  | 11110(4) | 1101(4)  |
| 5            |            |       |        | 1110(3) | 1110(3)  | 11110(4) | 1101(4)  | 1100(4)  |
| 6            |            |       |        |         | 11110(4) | 1101(4)  | 1100(4)  | 101(4)   |
| 7            |            |       |        |         |          | 1100(4)  | 101(4)   | 100(4)   |
| 8            |            |       |        |         |          |          | 100(4)   | 01(4)    |
| 9            |            |       |        |         |          |          |          | 00(4)    |

Рис. 10.4 иллюстрирует процедуру поиска оптимальной траектории для некоторого условного случая. Кружочками на рисунке отмечены узловые точки, в которых прерываются (или обрывается совсем) движение по данной траектории. Около каждой узловой точки в скобках указано соответствующее значение  $d_i$ ; римскими цифрами у узловых точек обозначен номер шага процедуры декодирования, на котором рассматривались варианты продолжения траектории от этой узловой точки. У ветвей кодового дерева проставлены

соответствующие значения информационных символов (единице отвечают верхние ветви). В табл. 10.1 приведены упорядоченные списки оценивавшихся вариантов траектории после каждого шага процедуры поиска для случая, изображенного на рис. 10.4 (в скобках указаны соответствующие значения  $d_i$ ). Каждый столбец этой таблицы характеризует информацию, хранящуюся в памяти после соответствующего шага процедуры (варианты траекторий обозначены соответствующими информационными последовательностями).

При осуществлении последовательного декодирования на основе посимвольного приема кодовой последовательности принятие решения о соответствующей ей информационной последовательности производится при исполнении использовании получаемой информации. Приписывая каждому принимаемому символу значение 0 или 1 ("жесткое" решение), при дальнейшей обработке мы учитываем лишь две градации значений отношения правдоподобия (меньше или больше порогового уровня). При некотором усложнении процедуры можно осуществить последовательное декодирование на основе приема кодовой последовательности в целом, обеспечивающее более достоверную оценку передаваемой информации. Усложнение процедуры сводится при этом к введению при приеме символов нескольких пороговых уровней отношения правдоподобия и последовательному сравнению с ними отношения правдоподобия, полученного для символа ("мягкое" решение). Если полученное отношение правдоподобия превышает верхний пороговый уровень, то символу приписывается значение 1, если нет, — то производится сравнение со следующим пороговым уровнем, при превышении которого приписывается следующая градация значения символа, и т.д. Если значения отношения правдоподобия меньше самого нижнего порогового уровня, символу приписывается значение 0. Например, при восьми пороговых уровнях каждому из принимаемых символов приписывается одно из следующих значений: 0, 1/8, 1/4, 3/8, 1/2, 5/8, 3/4, 7/8, 1. Отличие процедуры последовательного декодирования в случае использования мягкого решения при приеме символов состоит в обобщении понятия расстояния  $d_i$  между двоичными кодовыми последовательностями, отвечающими рассматриваемым траекториям, и принятой (уже не двоичной) последовательностью. Расстояние  $d_i$  должно определяться уже не простым счетом числа различающихся символов в сравниваемых последовательностях, а суммированием разностей значений этих символов (в общем случае величины  $d_i$  будут уже дробными). Усложнение процедуры последовательного декодирования сводится в основном к увеличению числа двоичных разрядов, требуемых для представления отдельных символов принятой последовательности и расстояний  $d_i$ , и к соответствующему увеличению объема памяти.

Вычислительная сложность алгоритмов последовательного декодирования определяется средним числом вычислительных операций на декодированный символ и дисперсией вычислительной нагрузки. Неравномерность вычислительной нагрузки, вытекающая из случайного характера длительностей траекторий, рассматриваемых на каждом шаге декодирования, является существенным недостатком последовательного декодирования. И дело здесь не просто в увеличении объема буферной памяти, требуемой для сглаживания вычислительной нагрузки с увеличением ее неравномерности, а в том, что и эта мера полностью не исключает вероятность появления такого пика вычислительной нагрузки, при котором буферная память будет переполнена. Переполнение же буферной памяти приводит к ошибкам декодирования. Именно этот источник ошибок в большинстве случаев определяет достоверность приема информации при последовательном декодировании. Для уменьшения вероятности ошибок, вызываемых переполнением буферной памяти, часто прибегают к таким мерам, которые с позиций теории кодирования понижают эффективность сверточного кода. К их числу относится, например, переход к систематическим кодам, позволяющим на время переполнения буферной памяти прекращать процедуру последовательного декодирования и выдавать информационные символы, занимающие определенные позиции в принятой кодовой последовательности, прямо на выход декодера без коррекции ошибок. Возможность искажения информационных символов при приеме кодовой последовательности при этом оказывается менее опасной, чем возможность появления ошибок, вызванных нарушением нормальной процедуры последовательного декодирования из-за переполнения буферной памяти.

К этой же категории мероприятий можно отнести разбиснение передаваемой кодовой последовательности на относительно короткие участки, разделенные последовательностью заранее известных символов (например, после определенного числа символов кодовой последовательности передается некоторое число единиц). Это позволяет начинать процедуру последовательного декодирования на каждом таком участке сначала, что снижает вероятность переполнения буферной памяти, а в случае ее переполнения на отдельных участках локализует вызываемое этим нарушение процедуры декодирования размером данного участка.

Совершенствование алгоритмов последовательного декодирования шло и ведется в направлении снижения среднего уровня вычислительной нагрузки и обеспечения большей ее равномерности. Существенное улучшение обоих этих показателей было достигнуто, в частности, с появлением алгоритма Зигангирова—Джеммика. В дальнейшем этот алгоритм был модифицирован Р. Фано с целью уменьшения требований к памяти за счет некоторого увеличения объема вычислений [31].

В специальной литературе, посвященной последовательному декодированию [28—32], можно найти большое число различных алгоритмов, отличающихся как частными особенностями процедуры декодирования, так и подходами к решению задачи. Большой интерес представляет, в частности предложенный А. Витерби алгоритм декодирования сверточных кодов, основанный на использовании усеченной функции правдоподобия [32]. Основным достоинством этого алгоритма, применимого лишь при небольшой длине кодовых ограничений  $k \leq 8$ , является фиксация вычислительной нагрузки на символ.

В последнее время все большее внимание привлекают так называемые каскадные коды, основанные на последовательном кодировании двумя разными кодами (Код, используемый при первом кодировании, называется внешним, а код используемый при втором кодировании, формирующий поступающие в линию символы, — внутренним). В частности, весьма эффективными оказываются каскадные коды, использующие в качестве внешнего соответствующие классы блочного кода, а в качестве внутреннего — сверточные коды [30].



**П.2. Матрицей** называется прямоугольная таблица чисел  $a_{ik}$  вида

$$[a_{ik}] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad (\text{П.9})$$

где индексы  $i$  и  $k$  определяют соответственно номера строки и столбца, на пересечении которых стоит данный элемент матрицы. Для обозначения матриц будем пользоваться прописной буквой:  $[a_{ik}] = A$ .

В общем случае, когда число строк не равно числу столбцов ( $m \neq n$ ), матрица называется *прямоугольной* размером  $m \times n$ . В частном случае, когда  $m = n$ , матрица называется *квадратной*.

Элементы матрицы  $a_{ik}$ ,  $k = 1, 2, \dots, n$ , составляющие  $i$ -ю строку матрицы, можно представить в виде компонент  $n$ -мерного вектора, называемого *вектор-строкой*. Аналогично элементы, составляющие  $k$ -й столбец матрицы, можно представить в виде компонент  $m$ -мерного вектор-столбца. Диагональ квадратной матрицы  $n \times n$ , проходящая из левого верхнего в правый нижний угол, т.е. состоящая из элементов  $a_{11}, a_{22}, \dots, a_{nn}$ , называется *главной диагональю*. Диагональ квадратной матрицы, проходящая из правого верхнего угла в левый нижний, т.е. составленная из элементов  $a_{ik}$  с суммой индексов  $i + k = n$ , называется *побочной диагональю*.

Квадратная матрица, у которой все элементы, расположенные вне главной диагонали, равны нулю, называется *диагональной*. Диагональная матрица с равными элементами главной диагонали  $d_{11} = d_{22} = \dots = d_{nn} = d$  называется *скалярной матрицей*. Диагональная матрица, у которой все элементы главной диагонали равны единице, называется *единичной*.

*Нулевой матрицей* называется матрица, все элементы которой равны нулю. В теории матриц ей отводится такая же роль, как нулю в теории чисел. Матрица называется *симметричной*, если ее элементы удовлетворяют условию  $a_{ik} = a_{ki}$ . *Транспонированной матрицей*  $A^T$  называется матрица, полученная из матрицы  $A$  превращением ее строк в столбцы матрицы  $A^T$  с тем же номером. Если матрица квадратная симметричная, то  $A = A^T$ .

*Суммой*  $A + B$  двух матриц  $A = [a_{ik}]$  и  $B = [b_{ik}]$  одинакового размера  $m \times n$  называется матрица  $C$  того же размера, каждый элемент которой равен сумме собственных элементов матриц  $A$  и  $B$ :

$$c_{ik} = a_{ik} + b_{ik}. \quad (\text{П.10})$$

Из (П.10) следует, что сложение матриц коммутативно.

*Произведением*  $\mu A$  матрицы  $A = [a_{ik}]$  на число  $\mu$  называется матрица  $A' = [a'_{ik}]$ , получающаяся умножением каждого элемента матрицы  $A$  на число  $\mu$ :

$$a'_{ik} = \mu a_{ik}. \quad (\text{П.11})$$

*Произведением*  $AB$  двух квадратных матриц  $A = [a_{ik}]$  и  $B = [b_{ik}]$  порядка  $n$  называется матрица  $C = [c_{ik}]$ , элемент которой  $c_{ik}$  равен скалярному произведению  $i$ -й вектор-строки матрицы-множимого на  $k$ -й вектор-столбец матрицы-множителя:

$$c_{ik} = \sum_{v=1}^n a_{iv} b_{vk}. \quad (\text{П.12})$$

Из определения следует, что операция умножения матриц некоммукативна (роли матриц в этой операции несимметричны: у первой берутся строки, у второй — столбцы). Не трудно убедиться, что если  $AB = C$ , то  $BA = C^T$ .

Из (П.12) следует, что если одна из матрицы-сомножителей, например матрица  $A$ , скалярная, т.е.

$$a_{iv} = \begin{cases} 0 & \text{при } v \neq i, \\ a & \text{при } v = i, \end{cases}$$

то

$$c_{ik} = \sum_{v=1}^n a_{iv} b_{vk} = a b_{ik},$$

т.е. умножение на скалярную матрицу сводится к умножению всех элементов матрицы-сомножителя на число, равное значению диагональных элементов скалярной матрицы. Таким образом, в этом случае операция умножения векторов вырождается в умножение вектора на скаляр (отсюда и название — скалярная матрица).

Роль единицы при умножении матриц играет единичная матрица  $E$ :

$$AE = A. \quad (\text{П.13})$$

Операция умножения может быть распространена и на прямоугольные матрицы при условии, что число столбцов матрицы-множимого равно числу строк матрицы-множителя. При этом матрица-произведение будет иметь такое же число строк как матрица-множимое и такое же число столбцов как матрица-множитель. Для прямоугольных матриц  $A$  и  $B$  одинакового размера  $m \times n$  может быть определена операция умножения на транспонированную матрицу  $AB^T$  или  $A^T B$ . При этом матрицы произведения будут квадратными размером  $m \times m$  и  $n \times n$  соответственно.

Матрица  $A$  и  $B$  одинакового размера  $m \times n$  называются *ортогональными*, если удовлетворяется условие

$$AB^T = 0, \quad (\text{П.14})$$

т.е. если соответствующие вектор-строки матриц  $A$  и  $B$  ортогональны.

Операции сложения и умножения матриц обладают свойством ассоциативности:

$$(A + B)C = AC + BC. \quad (\text{П.15})$$

С помощью матриц можно представить в компактной форме линейные преобразования многомерных величин (векторов). Пусть  $n$ -мерная величина, характеризуемая вектором  $X$  с компонентами  $x_1, x_2, \dots, x_n$ , подвергается линейному преобразованию вида



## Список литературы

1. Левин Б.Р. Теоретические основы статистической радиотехники. Книга первая. М., "Сов. радио", 1966.
2. Левин Б.Р. Теоретические основы статистической радиотехники. Книга вторая. М., "Сов. радио", 1968.
3. Тихонов В.И. Статистическая радиотехника. М., "Сов. радио", 1964.
4. Ван Трис Г. Теория обнаружения, оценок и модуляции. М., "Сов. радио", 1972.
5. Ширман Я.Д., Голиков В.Н. Основы теории обнаружения радиолокационных сигналов и измерения их параметров. М., "Сов. радио", 1963.
6. Трахтман А.М. Введение в обобщенную спектральную теорию сигналов. М., "Сов. радио", 1972.
7. Котельников В.А. Теория потенциальной помехоустойчивости. М., Госэнергоиздат, 1956.
8. Вудворд Ф.М. Теория вероятностей и теория информации с применением в радиолокации. Пер. с англ. М., "Сов. радио", 1955.
9. Миддлтон Д. Введение в статистическую теорию связи. Пер. с англ. Под ред. Б.Н. Левина. М., "Сов. радио", 1962.
10. Лезин Ю.С. Оптимальные фильтры и накопители импульсных сигналов. М., "Сов. радио", 1963.
11. Фэлькович С.Е. Оценка параметров сигнала. М., "Сов. радио", 1970.
12. Гуткин Л.С. Теория оптимальных методов радиоприема при флуктуационных помехах. М., "Сов. радио", 1972.
13. Вайнштейн Л.А., Зубков В.Д. Выделение сигналов на фоне случайных помех. М., "Сов. радио", 1960.
14. Винницкий А.С. Модулированные фильтры и следающий прием ЧМ сигналов. М., "Сов. радио", 1960.
15. Петрович Н.Т., Размахнин М.К. Системы связи с шумоподобными сигналами. М., "Сов. радио", 1969.
16. Вентцель Е.С. Теория вероятностей. М., Физматгиз, 1969.
17. Кловский Д.Д. Теория передачи сигналов. М., "Сов. радио", 1973.
18. Солодов А.В. Теория информации и ее применение к задачам автоматического управления и контроля. М., "Наука", 1967.
19. Харкевич А.А. Очерки общей теории связи. М., Гостехиздат, 1955.
20. Яглом А.М., Яглом И.М. Вероятность и информация. М., "Наука", 1971.
21. Гольдман С. Теория информации. Пер. с англ. М., ИЛ, 1957.
22. Теория информации и ее приложения. Под ред. А.А. Харкевича. М., Физматгиз, 1959.
23. Блох Э.Л., Харкевич А.А. К вопросу о геометрическом доказательстве теоремы Шеннона. — "Радиотехника", 1956, № 11.
24. Шеннон К. Работы по теории информации и кибернетике. М., ИЛ, 1963.
25. Харкевич А.А. Борьба с помехами. М., Физматгиз, 1963.
26. Питерсон У., Уэлдон Э. Коды, исправляющие ошибки. Пер. с англ. Под ред. Р.Л. Добрушина. М., "Мир", 1976.
27. Финк Л.М. Теория передачи дискретных сообщений. М., "Сов. радио", 1970.
28. Зигангиров К.Л. Процедуры последовательного декодирования. Статистическая теория связи, вып. 2. М., "Связь", 1974.
29. Возенкрафт Дж., Рейффен Б. Последовательное декодирование. Пер. с англ. Под ред. Р.Л. Добрушина. М., ИЛ, 1963.
30. Форми Д. Каскадные коды. Пер. с англ. под ред. С.И. Самойленко. М., "Мир", 1970.
31. Фано Р.М. Эвристическое рассмотрение метода статистического декодирования. — "Зарубежная радиоэлектроника", 1964, № 1.
32. Витерби А. Границы ошибок для сверточных кодов и асимптотически оптимальный алгоритм декодирования. — В кн.: Некоторые вопросы кодирования. Под ред. Э.Л. Блоха и М.С. Пинскера. М., ИЛ, 1970.
33. Финк Л.М. Теория передачи дискретных сообщений. М., "Сов. радио", 1970.
34. Бронштейн И.Н., Семендяев К.А. Справочник по математике. М., Физматгиз, 1959.
35. Ильин В.А., Поздняк Э.Г. Основы математического анализа. Ч. 1. М., "Наука", 1971.
36. Лозе М. Теория вероятностей. Пер. с англ. под ред. Ю.В. Прохорова. М., ИЛ, 1962.
37. Араманович И.Г., Берман А.Ф. Краткий курс математического анализа для вузов. М., "Наука", 1973.
38. В. Линдсей. Системы синхронизации в связи и управлении. Пер. с англ. Под ред. Ю.Н. Бакаева, М.В. Капранова. М., "Сов. радио", 1978.
39. Курош А.Г. Курс высшей алгебры. М., "Наука", 1965.
40. Гантмахер Ф.Р. Теория матриц. М., "Наука", 1967.
41. Карпелевич Ф.И., Садовский Л.Е. Элементы линейной алгебры и линейного программирования. Физматгиз, 1963.



## Предметный указатель

### Б

База сигнала 22

### В

Величины случайные некоррелированные 33  
-- ортогональные 33

### Д

Декодирование последовательное 197

### З

Закон биномиальный 29  
-- распределения дифференциальный 30  
-- интегральный 30  
-- нормальный 33  
---- многомерный 42

### И

Информация 127  
Источник сообщения 127

### К

Кодирование оптимальное статистическое 159  
Кодовое расстояние 164, 167  
Коды корректирующие 163  
-- блочные 163  
-- линейные 172  
-- непрерывные 164  
-- сверточные 165, 191  
-- несистематические 163, 193  
-- систематические 163, 192  
-- Хэмминга 176  
-- циклические 181  
Количество информации 127  
Критерий Байеса 55, 82  
-- идеального наблюдателя 60  
-- Неймана-Пирсона 68  
-- эффективности кодов 187

### М

Моменты распределения 30  
-- начальные 30  
-- центральные 31  
*m*-последовательность 104, 123, 186

### Н

Ненадежность канала 144

### О

Отношение правдоподобия 57, 86

### П

Полином генераторный 182  
-- порождающий 181  
Производительность источника сообщений 138  
Пропускная способность дискретного канала без помех 139  
---- с помехами 145  
-- непрерывного канала 158  
Процесс случайный 9  
-- нормальный 44  
-- стационарный в широком смысле 39  
-- строго стационарный 38  
-- эргодический 41

### Р

Равенство Парсеваля 17

### С

Сигнал детерминированный 9  
-- дискретный 10  
-- непрерывный 10  
-- случайный 9  
-- узкополосный 52  
-- цифровой 10  
-- широкополосный 98, 103  
Синдром проверки кодов 176

Скорость передачи информации 138

Сообщение 8, 127

Спектр сигнала 11

-- дискретный 11

-- комплексный 14

-- непрерывный 12

-- энергетический 47

### Т

Тело неопределенности 101

Теорема Котельникова 19

-- неопределенности 101

-- сложения вероятностей 26

-- умножения вероятностей 27

-- Хинчина-Винера 47

-- Шеннона для дискретного канала без помех 141

---- канала с помехами 147

-- непрерывного канала 173

### Ф

Фильтр идеальный 114

-- согласованный 119

Формула Байеса 29

Функции базисные 11, 109

-- собственные 109

Функция ковариационная 38

-- корреляционная 38

-- нормированная 40

-- неопределенности 101

-- передаточная 109

-- правдоподобия 86

-- риска 82

### Х

Характеристика импульсная переходная 110

### Ш

Шум белый 49

### Э

Энтропия источника дискретных сообщений 131

-- непрерывных сообщений 135

---- дифференциальная 136

## Оглавление

|  |          |
|--|----------|
| Предисловие .....  | 3        |
| <b>ЧАСТЬ I. ОСНОВЫ СТАТИСТИЧЕСКОЙ РАДИОТЕХНИКИ .....</b>   | <b>8</b> |
| Глава 1. <i>Общие сведения о сигналах</i> .....  | 8        |
| 1.1. Основные определения. Классификация сигналов .....  | 8        |
| 1.2. Формы представления сигналов .....  | 11       |
| 1.3. Дискретизация непрерывных сигналов. Теорема Котельникова .....  | 19       |
| Глава 2. <i>Случайные событие, величина, функция и их вероятностные характеристики</i> .....                           | 25       |
| 2.1. Случайное событие и вероятность .....   | 25       |
| 2.2. Случайные величины и их вероятностные характеристики .....  | 29       |
| 2.3. Нормальный закон распределения .....  | 33       |
| 2.4. Случайные функции и их вероятностное описание .....   | 36       |
| 2.5. Корреляционные характеристики случайных процессов. Стационарные и эргодические случайные процессы .....           | 37       |
| 2.6. Многомерный нормальный закон распределения. Нормальные случайные процессы .....                                   | 42       |
| 2.7. Спектральное представление стационарных случайных процессов .....   | 45       |
| 2.8. Случайные процесс типа белого шума .....  | 49       |
| 2.9. Узкополосные случайные сигналы .....  | 52       |
| Глава 3. <i>Статистическая теория обнаружения сигналов</i> .....   | 54       |
| 3.1. Классификация задач обнаружения сигналов .....  | 54       |
| 3.2. Общая постановка бинарной задачи обнаружения. Критерий Байеса .....   | 55       |
| 3.3. Синхронный прием символов двоичного цифрового кода при действии белого шума. Критерий идеального наблюдения ..... | 60       |
| 3.4. Радиолокационное обнаружение. Критерий Неймана – Пирсона .....  | 68       |
| 3.5. Многоальтернативные задачи обнаружения .....  | 75       |
| 3.6. Минимаксный критерий обнаружения сигналов .....   | 76       |
| 3.7. Особенности задачи селекции сигналов .....  | 79       |
| Глава 4. <i>Статистическая теория оценка параметров сигналов</i> .....   | 82       |
| 4.1. Байесовская оценка параметров сигнала. Функция риска .....  | 82       |
| 4.2. Оптимальная оценка параметров сигнала при действии нормального белого шума .....                                  | 86       |

|   |            |
|---|------------|
| 4.3. Особенности оптимальной оценки параметров сигнала, зависящего от других случайных параметров, не подлежащих оценке ..... | 91         |
| 4.4. Потенциальная точность измерения временной задержки и доплеровского сдвига частоты сигнала .....                         | 94         |
| 4.5. Совместная оценка фазы и частоты сигнала. Теорема неопределенности .....   | 98         |
| 4.6. Широкополосные сигналы .....   | 103        |
| Глава 5. <i>Фильтрация случайных сигналов</i> .....   | 107        |
| 5.1. Общая характеристика задач фильтрации .....  | 107        |
| 5.2. Краткие сведения о линейных цепях и линейных преобразованиях сигналов .....  | 109        |
| 5.3. Оптимальная линейная фильтрация по критерию минимума искажений полезного сигнала .....                                   | 113        |
| 5.4. Оптимальная линейная фильтрация по критерию максимума отношения сигнал-шум (согласованные фильтры) .....                 | 117        |
| 5.5. Примеры согласованных фильтров .....   | 121        |
| <b>ЧАСТЬ II. ОСНОВЫ ТЕОРИИ ИНФОРМАЦИИ И ИДЕАЛЬНОЕ КОДИРОВАНИЕ .....</b>   | <b>127</b> |
| Глава 6. <i>Количественное определение информации и информационные характеристики источников сообщений</i> .....              | 127        |
| 6.1. Количественное определение информации, содержащейся в сообщении .....  | 127        |
| 6.2. Информационные характеристики источников дискретных сообщений .....  | 130        |
| 6.3. Энтропия источников непрерывных сообщений .....  | 135        |
| Глава 7. <i>Пропускная способность каналов связи и теоремы оптимального кодирования</i> .....                                 | 138        |
| 7.1. Пропускная способность дискретного канала без помех. Теорема Шеннона для канала без помех .....                          | 138        |
| 7.2. Скорость передачи информации и пропускная способность дискретного канала при наличии помех .....                         | 143        |
| 7.3. Основная теорема Шеннона для дискретного канала с помехами .....   | 146        |
| 7.4. Пропускная способность непрерывного канала при наличии аддитивного шума .....  | 152        |
| <b>ЧАСТЬ III. ОСНОВЫ ТЕОРИИ КОДИРОВАНИЯ .....</b>   | <b>155</b> |
| Глава 8. <i>Общие сведения о кодировании</i> .....  | 155        |
| 8.1. Функциональная схема радиолинии передачи дискретных сообщений .....  | 155        |
| 8.2. Цифровые коды .....  | 157        |
| 8.3. Основные задачи теории кодирования .....   | 157        |
| 8.4. Понятие об экономном кодировании .....   | 158        |
| 8.5. Общая характеристика задачи помехоустойчивого кодирования .....  | 162        |

|  |     |
|--|-----|
| Глава 9. Блочные корректирующие коды .....                                   | 166 |
| 9.1. Основные характеристики и корректирующие свойства блочных кодов .....   | 166 |
| 9.2. Линейные корректирующие коды. Коды Хэмминга .....                       | 172 |
| 9.3. Циклические коды .....  | 181 |
| 9.4. Критерии эффективности блочных корректирующих кодов .....               | 187 |
| Глава 10. Сверточные корректирующие коды .....                               | 191 |
| 10.1. Особенности кодирования и декодирования сверточных кодов .....         | 191 |
| 10.2. Сверточные коды с синдромной коррекцией .....                          | 195 |
| 10.3. Сверточные коды с использованием последовательного декодирования ..... | 197 |
| Приложение. Краткие сведения об определителях и матрицах .....               | 202 |
| Список литературы .....  | 208 |
| Предметный указатель .....   | 210 |

**Липкин Исаак Аронович**  
**СТАТИСТИЧЕСКАЯ РАДИОТЕХНИКА**  
**ТЕОРИЯ ИНФОРМАЦИИ И КОДИРОВАНИЯ**

Под редакцией автора

Компьютерная верстка *А.А. Гурова*

Подписано в печать 15.02.2002.  
 Печать офсетная. Формат 60 х 84/16  
 Печ. л. 13,5. Тираж 1000 экз.

ЗАО «Издательское предприятие «Вузовская книга»  
 125871, Москва, Волоколамское шоссе, д. 4  
 Тел. 158-02-35  
 E-mail: vbook@mai.ru