

ТЕОРИЯ ПЕРЕДАЧИ ЭЛЕКТРИЧЕСКИХ СИГНАЛОВ ПРИ НАЛИЧИИ ПОМЕХ

СБОРНИК ПЕРЕВОДОВ

Под редакцией

Н. А. ЖЕЛЕЗНОВА

УПРАВЛЕНИЕ НАУЧНОЙ ИНФОРМАЦИИ

ИЗДАТЕЛЬСТВО
ИНОСТРАННОЙ ЛИТЕРАТУРЫ

Москва—1953

ПРЕДИСЛОВИЕ

Современный этап развития радиотехники характеризуется широким применением методов теории вероятностей для решения многих радиотехнических задач. В связи с этим представляет интерес попытка создания общей статистической теории передачи электрических сигналов в системах связи, радиолокации, телемеханики при наличии помех.

Характерной особенностью этой теории является то, что сигналы на входе канала рассматриваются не как заданные функции времени, а как множество возможных функций времени, определенных вместе с вероятностями их появления. Другой отличительной чертой теории является то, что воздействие помех оценивается не по отношению $\frac{\text{сигнал}}{\text{помеха}}$ на выходе канала, а по некоторому статистическому параметру, характеризующему достоверность полученных данных. Так, например, процесс радиолокационного измерения дальности оценивается ненадежностью измерения, под которой понимается вероятность того, что во множестве результатов наблюдения измеренное значение дальности не лежит вблизи истинного значения.

Основные результаты теории сводятся к установлению предельных соотношений для пропускной способности канала, ограниченной отношением интенсивностей сигнала и помех, а также полосой радиочастот, занимаемой каналом. Утверждается, что предельная пропускная способность может быть достигнута при сколь угодно малой частоте ошибок в воспроизведении переданных сигналов. Это обеспечивается выбором надлежащего метода кодирования передаваемых сообщений.

Таким образом, открывается возможность оценки различных систем передачи электрических сигналов при наличии помех путем сравнения действительно достигаемой пропускной способности с предельно возможной.

Предлагаемый вниманию читателя сборник состоит из трех частей. Первая часть посвящена статистической теории передачи электрических сигналов при наличии помех различных типов. Рассматривается как система с дискретными сигналами, так и система с непрерывными сигналами.

Во второй части излагается теория внутренних помех радиоканала — флуктуационных «шумов», ограничивающих предельную пропускную способность радиоканала. Эти материалы не только необходимы для изучения I и III частей сборника, но и представляют также самостоятельный интерес.

Третья часть сборника служит иллюстрацией применения статистических методов к практическим задачам радиолокации — исследованию точности и достоверности измерений дальности до стационарной цели, проектированию радиолокационных приемников на основе статистических методов и т. д.

Следует отметить, что в развитии статистических методов исследования процессов передачи и приема электрических сигналов при наличии помех вклад нашей отечественной науки исключительно велик. Необходимый для этого математический аппарат был создан трудами знаменитых русских математиков П. Л. Чебышева, А. А. Маркова, А. М. Ляпунова, трудами советских ученых А. Н. Колмогорова, А. Я. Хинчина, С. Н. Бернштейна и др. Значение этих работ настолько велико, что без них нельзя себе представить современной теории вероятностей.

При изучении системы с дискретными сигналами используются случайные процессы с прерывным временем (цепи Маркова), названные по имени создателя теории цепей А. А. Маркова.

Исключительное значение в радиотехнике имеют стационарные случайные процессы, теория которых была разработана А. Я. Хинчиным и А. Н. Колмогоровым. В принадлежащей А. Я. Хинчину теории корреляции стационарных случайных процессов содержится мощный математический аппарат, широко используемый при решении разнообразных радиотехнических задач. Большое значение имеет, например, теорема о связи между спектральной плотностью случайного процесса и его функцией корреляции, лежащая в основе теории внутренних помех радиоканала.

Применение статистических методов к задачам помехоустойчивости радиоприема было заложено работами В. И. Сифорова, выполненными в середине 30-х годов.

Проблема пропускной способности канала связи была поставлена В. А. Котельниковым еще в 1933 г. Тогда же В. А. Котельников доказал теорему о том, что непрерывная функция времени с ограниченным спектром полностью определяется заданием ряда чисел, связанных со значениями этой функции в дискретные моменты времени. Эта замечательная теорема имеет фундаментальное значение для современной теории передачи электрических сигналов. Следует заметить, что в иностранной технической литературе она была сформулирована значительно позднее.

Вопрос о пропускной способности канала получил дальнейшее развитие в работах Д. В. Агеева (1938 г.), указавшего на существенную роль помех в ограничении пропускной способности.

Идеи о геометрической трактовке процессов передачи и приема сигналов при наличии в канале внутренних шумов, широко используемые в статистической теории связи, также были впервые высказаны и применены для рассмотрения практических вопросов В. А. Котельниковым еще в 1946 г.

Большое значение для развития теории внутренних помех в радиоканале имела монография В. Л. Грановского «Электрические флуктуации», вышедшая в свет в 1936 г.

Существенным вкладом в теорию флуктуационных шумов явились труды М. А. Леонтовича, В. И. Бунимовича и других советских ученых. В работах В. И. Бунимовича, начатых в 1940 г., получил развитие метод исследования преобразования сигнала и шумов нелинейной системой.

Нужно сказать, что статистическая теория передачи электрических сигналов при наличии помех находится еще на начальном этапе своего развития, однако это новое направление в современной радиотехнике представляет безусловный интерес. Можно не сомневаться, что статистические методы сыграют важную роль в решении основной проблемы современной радиотехники — разработке способов передачи и приема сигналов, обеспечивающих большую помехоустойчивость, чем существующие методы.

Следует отметить, что включенные в этот сборник работы иностранных авторов не лишены недостатков.

Прежде всего нужно указать на тенденциозное замалчивание работ советских авторов, значение которых для развития статистических методов в радиотехнике исключительно велико. Только у Райса имеется ссылка на статью А. Я. Хинчина по теории корреляции стационарных случайных процессов, другие труды советских ученых обойдены молчанием.

В работе Шэннона проявляется стремление автора придать изложению формальный характер, сторванный от практических задач. Это в известной мере затрудняет усвоение материала и применение его на практике.

Терминология Шэннона тоже в ряде случаев вызывает возражения. Так, автор на основании чисто внешнего сходства математической формулы называет величину H , через которую определяется пропускная способность канала, «энтропией». Таким образом, понятие энтропии отрывается от конкретного физического содержания, которое в него вкладывается в статистической физике. Это открывает дорогу различным формалистическим построениям, чем действительно и занялись вскоре после опубликования работы Шэннона некоторые иностранные авторы.

В работах Вудворда и Девиса по применению статистических методов в радиолокации проявляется субъективное понимание математической вероятности. Некоторые рассуждения в тексте I части свидетельствуют о том, что авторы иногда пытаются

трактовать вероятность как величину, характеризующую «степень уверенности» наблюдателя.

Искусственность такого подхода очевидна из самого характера рассматриваемого физического процесса радиолокационного измерения дальности при наличии помех. Распределение апостериорных вероятностей для различных значений дальности до цели после приема сигнала характеризует объективный характер связи между приемом сигнала и результатами измерения дальности, а вовсе не субъективные впечатления наблюдателя. Это подтверждается хотя бы тем, что можно построить систему измерения дальности до цели, которая на основании расчета распределения апостериорных вероятностей выдает наивероятнейшее значение дальности, полностью исключая субъективную оценку наблюдателя.

Необходимо отметить, что терминология статистической теории передачи электрических сигналов и ряд ее понятий используются некоторыми зарубежными математиками и инженерами для спекуляций, связанных с пресловутой «кибернетикой».

Так, например, Н. Винер, С. Гольдман и др., исходя из внешней, поверхностной аналогии и спекулируя на нечеткости и двусмысленности некоторых терминов и понятий, пытаются перенести закономерности радиосвязи на биологические и психологические явления, говорят о «пропускной способности» человеческого мозга и т. д. Естественно, что все эти попытки придать кибернетике наукообразный характер с помощью заимствованных из другой области терминов и понятий отнюдь не делают кибернетику наукой — она остается лжетеорией, созданной реакционерами от науки и философствующими невеждами, находящимися в плену идеализма и метафизики. В то же время досужие упражнения философствующих лжеученых набрасывают тень на статистическую теорию передачи сигналов при наличии помех, результаты и выводы которой сами по себе имеют большое научное и прикладное значение.

Это обстоятельство пришлось учесть при подготовке к печати настоящего сборника. При редактировании сборника были произведены небольшие сокращения, причем текст был освобожден от некоторых ненаучных идеалистических построений, никак не связанных с основным материалом. Ряд неудачных и двусмысленных терминов был заменен с целью устранения возможности неправильного толкования некоторых понятий.

Из текста статей исключены исторические справки, носящие тенденциозный и поверхностный характер. Ссылки авторов на книги, имеющиеся в русском переводе, заменены ссылками на соответствующие советские издания. Кроме того, сборник дополнен кратким указателем литературы, составленным редактором.

СТАТИСТИЧЕСКАЯ ТЕОРИЯ ПЕРЕДАЧИ ЭЛЕКТРИЧЕСКИХ СИГНАЛОВ¹⁾

К. ШЭННОН

ВВЕДЕНИЕ

Развитие различных методов модуляции, таких, как импульсно-кодовая и временная импульсная модуляции, которые уменьшают требования к отношению сигнал/шум за счет расширения полосы частот, повысило интерес к общей теории связи²⁾.

В настоящей работе мы расширим теорию с тем, чтобы включить в нее некоторое число новых факторов, в частности влияние шумов в канале и возможность улучшения связи за счет использования статистической структуры исходного сообщения и свойств оконечного получателя сообщений.

Основная задача связи заключается в точном или приближенном воспроизведении в одной точке сообщения, выбранного в некоторой другой точке. Существенно, что действительное сообщение является одним, выбранным из определенного множества возможных сообщений. Система должна быть спроектирована таким образом, чтобы она обеспечивала передачу любого возможного сообщения, а не только того, которое действительно будет выбрано, так как последнее в момент проектирования еще не известно.

Если число сообщений во множестве конечно, то это число или некоторая монотонная функция от него может быть принята за меру количества данных³⁾, передаваемых тогда, когда из множества выбирается одно сообщение, причем все возможности выбора равновероятны.

¹⁾ C. E. Shannon and W. Weaver «The Mathematical Theory of Communication», The University of Illinois Press, 3—89, 1949.

²⁾ Термины «связь» и «система связи» понимаются автором весьма широко. Системой связи может быть любая система, предназначенная для передачи и приема сигналов, будь то система радиосвязи, радиолокации, телемеханики и пр. (Прим. ред.)

³⁾ В оригинале применяется термин «информация». Поскольку, однако, автор в дальнейшем придает ему специальное значение, устраняя семантические аспекты этого термина, мы от него отказались. (Прим. ред.)

Как было указано Хартлеем¹⁾, наиболее естественно выбрать логарифмическую функцию. Хотя это определение должно быть существенно обобщено при рассмотрении статистических свойств сообщений, а также при наличии непрерывной совокупности сообщений, мы будем во всех случаях пользоваться логарифмической мерой.

Логарифмическая мера обладает рядом существенных преимуществ.

1. Она практически наиболее удобна. Параметры, имеющие техническое значение, как, например, время, ширина полосы частот, число реле и т. п., имеют тенденцию изменяться линейно с логарифмом числа возможностей. Например, добавление одного реле к существующей группе удваивает число возможных положений реле. Это прибавляет единицу к логарифму этого числа при основании 2. Удвоение времени, грубо говоря, возводит число возможных сообщений в квадрат, т. е. удваивает логарифм и т. д.

2. Она ближе к нашему интуитивному представлению о подходящей мере. Это обстоятельство тесно связано с первым, так как мы интуитивно измеряем величины путем линейного сравнения с принятыми эталонами. Каждый, например, чувствует, что две перфорированные карточки содержат вдвое больший запас сведений, а два идентичных канала имеют удвоенную пропускную способность.

3. Она более удобна с математической точки зрения. Многие предельные переходы весьма просты при применении логарифмов, но потребовали бы сложных приемов при использовании самого числа возможностей.

Выбор основания логарифмов соответствует выбору единицы измерения количества данных. При основании 2 получаются единицы, которые могут быть названы «двоичными единицами». Прибор с двумя состояниями равновесия, например реле или спусковая схема, может запасти одну двоичную единицу. N таких приборов могут запасти N единиц, так как общее число возможных состояний равно 2^N и $\log_2 2^N = N$. Если же выбрано основание 10, то единицы могут быть названы «десятичными единицами». Так как

$$\log_2 M = \frac{\log_{10} M}{\log_{10} 2} = 3,32 \log_{10} M,$$

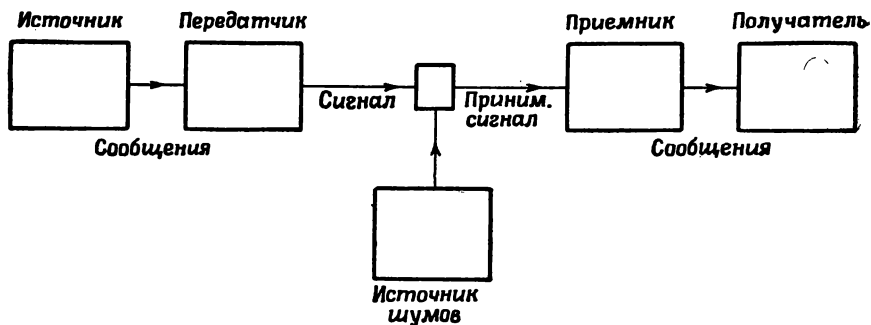
то десятичная единица равна примерно $3\frac{1}{3}$ двоичных единиц. Числовое колесо арифмометра имеет десять положений равновесия и поэтому обладает способностью запасти одну десятичную единицу. При аналитических расчетах, когда приходится интегрировать и дифференцировать, иногда удобно применять основание e . Получающиеся при этом единицы могут быть названы «натуральными

¹⁾ *BSTJ*, 535, July 1928.

единицами». Переход от основания a к основанию b требует лишь умножения на $\log_b a$.

В дальнейшем под системой связи будем подразумевать систему, схематически показанную на фиг. 1. Она состоит из пяти основных частей:

1. *Источник*, создающий сообщение или последовательность сообщений, которые должны быть переданы на приемный конец. Сообщение может быть различного типа, например: а) последовательность букв, как в системах телеграфа или телетайпа; б) одиночная функция времени $f(t)$, как в телефонии или радиотелефонии;



Фиг. 1. Схема общей системы связи.

в) функция времени и других переменных, как в черно-белом телевидении, где сообщение можно рассматривать как функцию $f(x, y, t)$ от двух пространственных координат и времени (интенсивность света в точке (x, y) в момент t на экране приемной трубки); г) две (или более) функции времени, скажем, $f(t)$, $g(t)$, $h(t)$, как это бывает в случае передачи «трехмерного» звука, или если система рассчитана на обслуживание нескольких индивидуальных каналов по многократному способу; д) несколько функций от нескольких переменных — в цветном телевидении сообщение состоит из трех функций $f(x, y, t)$, $g(x, y, t)$, $h(x, y, t)$, определенных в трехмерном континууме. Мы можем рассматривать эти три функции как составляющие векторного поля в некоторой области; подобно этому несколько источников черно-белого телевидения будут давать сообщения, состоящие из нескольких функций от трех переменных; е) различные комбинации перечисленных случаев, например, в телевидении с каналом звукового сопровождения.

2. *Передатчик*, который преобразует сообщение таким образом, что получается сигнал, пригодный для передачи по каналу. В телефонии эта операция состоит просто в преобразовании звукового давления в пропорционально изменяющийся электрический ток.

В телеграфии имеется операция кодирования, которая дает последовательность точек, тире и пробелов, соответствующих сообщению. В системах многоканальной кодовой импульсной модуляции значения различных речевых функций должны быть зафиксированы в некоторые дискретные моменты времени, компрессированы, квантованы по амплитудам, закодированы и, наконец, соответствующим образом смешаны для образования сигнала. Системы с вокодером¹⁾, телевидение и частотная модуляция являются другими примерами сложных операций — преобразования сообщения в сигнал.

3. *Канал* — среда, используемая для передачи сигнала от передатчика к приемнику. Это может быть пара проводов, коаксиальный кабель, полоса радиочастот, луч света и т. д.

4. *Приемник* обычно выполняет операции, обратные осуществленным в передатчике, восстанавливая сообщение из сигнала.

5. *Получатель* — это лицо (или аппарат), для которого предназначено сообщение.

В дальнейшем будет рассмотрено несколько общих проблем, относящихся к системам связи. Для этого прежде всего необходимо описать различные элементы при помощи математических величин, должным образом идеализированных по сравнению со своими физическими оригиналами.

Системы связи можно грубо подразделить на три главные категории: дискретные, непрерывные и смешанные. Под дискретной системой будем понимать систему, в которой как сообщение, так и сигнал представляют собой последовательность дискретных символов²⁾. Типичным случаем является телеграфия, где сообщение — последовательность букв, а сигнал — последовательность точек, тире и пробелов. В непрерывной системе как сигнал, так и сообщение рассматриваются как непрерывные функции, например в радиотелефонии и телевидении. В смешанных системах имеются как непрерывные, так и дискретные переменные, примером чего может служить передача речи посредством кодовой импульсной модуляции.

Рассмотрим сначала дискретный случай. Он имеет применение не только в теории связи, но также в теории счетных машин и в других областях. Кроме того, дискретный случай является основой для рассмотрения непрерывного и смешанного случаев, которые исследуются во второй половине работы.

¹⁾ Вокодер — устройство «синтетической телефонии», в которой подлежащие передаче сигналы синтезируются на приемном конце. По каналу передаются только «командные сигналы», получаемые в результате анализа звуков передаваемой естественной речи. Эти сигналы управляют процессом синтеза (высотой, силой тонов, ритмом речи и т. д.) (*Прим. ред.*)

²⁾ Символами здесь и в дальнейшем автор называет элементы сообщения или сигнала, например буквы, точки, тире и т. д. (*Прим. ред.*)

Г л а в а I

ДИСКРЕТНЫЕ СИСТЕМЫ БЕЗ ШУМОВ

1. ДИСКРЕТНЫЙ КАНАЛ БЕЗ ШУМОВ

Телеграф и телетайп являются двумя простейшими примерами дискретного канала. Вообще же дискретный канал означает систему, в которой может быть передана из одной точки в другую последовательность наборов из конечного ряда элементарных символов S_1, \dots, S_n . Предполагается, что символ S_i имеет длительность во времени t_i секунд (не обязательно, чтобы все символы обладали одинаковой длительностью, например точки и тире в телеграфии). Не обязательно также, чтобы все возможные последовательности символов S_i могли передаваться системой; могут допускаться только некоторые последовательности. Это будут возможные сигналы для канала. Например, в телеграфии предполагаются следующие символы:

1. Точка, состоящая из замыкания линии на некоторую единицу времени и последующего размыкания на такое же время.

2. Тире, состоящее из замыкания на три единицы времени и размыкания на одну единицу.

3. Пробел между буквами, состоящий, скажем, из размыкания на три единицы.

4. Пробел между словами — размыкание линии на шесть единиц времени.

Необходимо наложить ограничения на допустимые последовательности, чтобы пробелы не следовали друг за другом, так как два промежутка между буквами дают промежуток между словами.

Теперь рассмотрим вопрос о том, каким образом можно измерить пропускную способность такого канала.

В случае телетайпа, где все символы обладают одинаковой длительностью и допустимы все последовательности из 32 символов, ответ очень прост. Каждый символ представляет собой пять двоичных единиц. Если система передает n символов в 1 сек., естественно сказать, что канал обладает способностью передачи в 5 n двоичных единиц в секунду. Это не означает, что канал телетайпа будет всегда передавать сообщения с такой скоростью. Это — максимальная возможная скорость, и будет ли в действительности достигнут этот максимум, зависит от источника сообщений на входе канала.

В более общем случае символов различной длительности и при ограничении допустимых последовательностей можно дать следующее определение:

О п р е д е л е н и е: Пропускная способность C дискретного канала выражается формулой

$$C = \lim_{T \rightarrow \infty} \frac{\log N(T)}{T},$$

где $N(T)$ — число допустимых сигналов длительностью T .

Легко видеть, что в случае телетайпа это определение сводится к предыдущему. Можно показать, что рассматриваемый предел существует как конечное число в большинстве представляющих интерес случаев. Предположим, что допустимы все последовательности символов S_1, S_2, \dots, S_n и что эти символы имеют длительности t_1, t_2, \dots, t_n . Какова тогда пропускная способность канала? Если $N(t)$ означает число последовательностей длительностью t , то

$$N(t) = N(t - t_1) + N(t - t_2) + \dots + N(t - t_n),$$

т. е. общее число равно сумме чисел последовательностей, оканчивающихся символами S_1, S_2, \dots, S_n , а эти числа соответственно равны $N(t - t_1), N(t - t_2), \dots, N(t - t_n)$.

Согласно хорошо известному положению исчисления конечных разностей $N(t)$ при больших t асимптотически приближается к AX_0^t , где A — постоянная, а X_0 — наибольший вещественный корень характеристического уравнения

$$X^{-t_1} + X^{-t_2} + \dots + X^{-t_n} = 1$$

и, следовательно:

$$C = \lim_{T \rightarrow \infty} \frac{\log AX_0^T}{T} = \lg X_0.$$

При наличии ограничений, наложенных на допустимые последовательности, часто все же можно получить уравнение в конечных разностях того же типа и найти C из характеристического уравнения. В упомянутом случае телеграфии

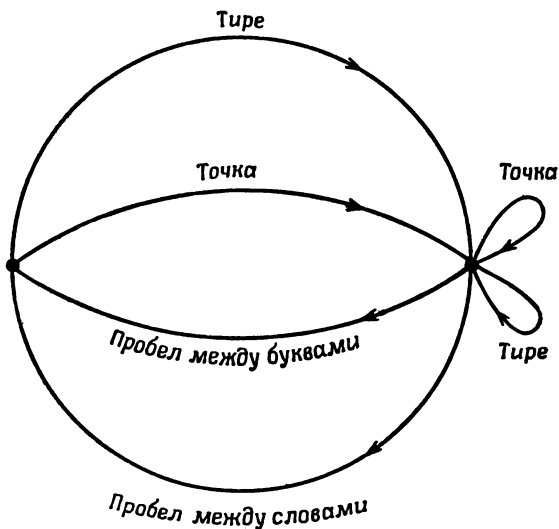
$$\begin{aligned} N(t) = & N(t-2) + N(t-4) + N(t-5) + N(t-7) + \\ & + N(t-8) + N(t-10), \end{aligned}$$

в чем можно убедиться, подсчитывая последовательности символов с учетом последнего или следующего за последним символа. Отсюда C есть $-\log \mu_0$, где μ_0 — положительный корень уравнения

$$1 = \mu^2 + \mu^4 + \mu^5 + \mu^7 + \mu^8 + \mu^{10}.$$

Решая это уравнение, найдем $C = 0,539$.

Весьма общий вид ограничений, которые могут быть наложены на допустимые последовательности, состоит в следующем. Вообразим некоторое число возможных состояний a_1, a_2, \dots, a_m . В каждом состоянии могут быть переданы только некоторые символы из ряда S_1, \dots, S_n (различные наборы для разных состояний). Когда один из этих символов передан, состояние переходит в некоторое новое состояние в зависимости как от старого состояния, так и от переданного символа. Простейшим примером этого является телеграфия. Имеются два состояния в зависимости от того, был ли последним



Фиг. 2. Графическое представление ограничений, наложенных на телеграфные символы.

переданным символом пробел или нет. Если был пробел, то после этого могут быть переданы только точка или тире, и состояние непременно изменится. Если нет, то может быть передан любой символ и состояние либо изменится, если передан пробел, либо в противном случае не изменится.

Все это может быть иллюстрировано графиком, показанным на фиг. 2. Точки разветвления соответствуют состояниям, а линии указывают символы, возможные в данном состоянии, а также результирующие состояния. В Приложении 1 показывается, что если условие относительно допустимых последовательностей может быть задано в такой форме, то существует определенное значение пропускной способности, которое может быть вычислено согласно следующей теореме.

Теорема 1

Пусть $b_{ij}^{(s)}$ означает длительность s -го символа, который возможен в состоянии i и ведет к состоянию j . Тогда пропускная способность канала C равна $\log W$, где W — наибольший вещественный корень уравнения в виде определителя

$$\left| \sum_s W^{-b_{ij}^{(s)}} - \delta_{ij} \right| = 0,$$

где $\delta_{ij} = 1$, если $i = j$, и нуль в противном случае.

Например, в случае телеграфии определитель имеет вид

$$\begin{vmatrix} -1 & (W^{-2} + W^{-4}) \\ (W^{-3} + W^{-6}) & (W^{-2} + W^{-4} - 1) \end{vmatrix} = 0.$$

Разложение этого определителя дает уравнение, которое приводилось выше для этого случая.

2. ИСТОЧНИК ДИСКРЕТНЫХ СООБЩЕНИЙ

Мы уже видели, что при весьма общих условиях логарифм числа возможных сигналов в дискретном канале линейно возрастает со временем. Пропускная способность может быть определена указанием скорости этого возрастания: числа двоичных единиц в секунду, требуемого для задания отдельного применяемого сигнала.

Рассмотрим теперь источник сообщений. Как следует математически описывать источник и какое количество данных, измеренное в двоичных единицах в секунду, создает такой источник? Знание статистических свойств источников имеет большое значение для уменьшения необходимой пропускной способности канала путем рационального кодирования сообщений. Например, в телеграфии передаваемые сообщения состоят из последовательностей букв. Эти последовательности, однако, не вполне хаотичны. Вообще говоря, они образуют фразы и имеют статистическую структуру, скажем, английского языка. Буква E появляется много чаще, чем Q , последовательность TH чаще, чем XP , и т. д. Наличие такой структуры позволяет экономить время (или пропускную способность канала) путем рационального кодирования последовательностей сообщений в последовательности сигналов.

В ограниченных пределах это всегда делается в телеграфии: самый короткий символ в канале — точка применяется для наиболее частой в английском языке буквы E , в то время как редкие буквы Q , X , Z выражаются более длинными последовательностями тире и точек. Еще отчетливее проводится этот принцип в некоторых коммерческих кодах, где наиболее обычные слова и фразы изображаются кодовыми группами из четырех или пяти букв, что дает значительную экономию среднего времени.

Можно себе представить, что дискретный источник создает сообщение символ за символом. Последовательные символы выбираются соответственно некоторым вероятностям, зависящим, вообще говоря, как от предыдущего выбора, так и от данного рассматриваемого символа, о котором идет речь.

Математическая модель системы, которая создает такую последовательность символов, управляемую совокупностью вероятностей, известна под названием стохастического процесса. Поэтому можно считать, что дискретный источник может быть представлен некоторым стохастическим процессом. Обратно, любой стохастический процесс, который дает дискретную последовательность символов, выбираемых из конечного ряда, может рассматриваться как описывающий некоторый дискретный источник. Это включает такие случаи, как:

1. Источники непрерывных сообщений, которые превращены в дискретные путем квантования. Например, квантованная речь от передатчика с кодовой импульсной модуляцией или квантованный телевизионный сигнал.

2. Математические случаи, когда просто абстрактно определяется некоторый стохастический процесс, создающий последовательность символов. Приведем примеры источников последнего типа.

А. Пусть имеются пять букв A, B, C, D, E , которые выбираются с вероятностью 0,2 каждая, независимо от предыдущей буквы. Это привело бы к последовательности такого примерно вида:

*B D C B C E C C C A D C B D D A E C E E A
A B B D A E E C A C E E B A E E C B C E A D*

Этот пример был построен при помощи таблицы случайных чисел.

Б. Используются те же пять букв, но с вероятностями 0,4, 0,1, 0,2, 0,2, 0,1 соответственно. Следующие друг за другом буквы выбираются независимо от предыдущих. Типичное сообщение от такого источника имеет вид

*A A A C D C B D C E A A D A D A C E D A
E A D C A B E D A D D C E C A A A A A D*

В. Более сложная структура получается, если последующие символы не выбираются независимо, так что их вероятности зависят от предшествующих букв. В простейшем случае выбор зависит только от непосредственно предшествующей буквы, а не от ранее стоящих букв. Тогда статистическая структура может быть описана набором вероятностей перехода $p_i(j)$, т. е. вероятностей того, что за буквой i следует буква j . Индексы i и j охватывают все возможные символы. Другой эквивалентный способ описания структуры состоит в задании вероятностей двухбуквенных сочета-

ний $p(i, j)$, т. е. относительных частот двухбуквенного сочетания (i, j) . Частоты появления букв $p(i)$ (вероятность буквы i), вероятности переходов $p_i(j)$ и вероятности двухбуквенных сочетаний $p(i, j)$ связаны следующими соотношениями:

$$p(i) = \sum_j p(i, j) = \sum_j p(j, i) = \sum_j p(j) p_j(i),$$

$$p(i, j) = p(i) p_i(j),$$

$$\sum_j p_i(j) = \sum_i p(i) = \sum_{i,j} p(i, j) = 1.$$

В качестве частного примера предположим, что имеются три буквы с таблицами вероятностей

$p_i(j)$	A	B	C
A	0	$\frac{4}{5}$	$\frac{1}{5}$
B	$\frac{1}{2}$	$\frac{1}{2}$	0
C	$\frac{1}{2}$	$\frac{2}{5}$	$\frac{1}{10}$

i	$p(i)$
A	$\frac{9}{27}$
B	$\frac{16}{27}$
C	$\frac{2}{27}$

$p(i, j)$	A	B	C
A	0	$\frac{4}{15}$	$\frac{1}{15}$
B	$\frac{8}{27}$	$\frac{8}{27}$	0
C	$\frac{1}{27}$	$\frac{4}{135}$	$\frac{1}{135}$

Типичное сообщение от этого источника имеет вид

A B B A B A B A B A B A B A B B V A B V
 B B B A B A B A B A B A B V V A C A C A
 B B A B V B V A B V A C V V B A B A

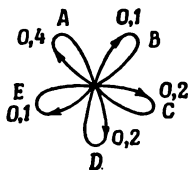
Следующее повышение сложности состоит в учете частот появления трехбуквенных сочетаний, но не более. Выбор буквы будет зависеть от предшествующих двух букв, но не от того, что было до этих букв. При этом должна быть задана совокупность частот трехбуквенных сочетаний $p(i, j, k)$ или эквивалентная совокупность вероятности переходов $p_{ij}(k)$. Продолжая таким образом, можно последовательно получать все более сложные стохастические процессы. В общем случае сочетаний из n букв для определения статистической структуры требуется совокупность вероятностей n -буквенных сочетаний $p(i_1, i_2, \dots, i_n)$ или вероятностей переходов $p_{i_1, i_2, \dots, i_{n-1}}(i_n)$.

3. ГРАФИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ ЦЕПЕЙ МАРКОВА

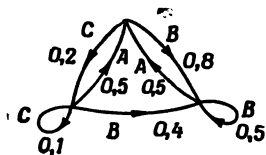
Стохастический процесс описанного типа известен в математике как дискретная цепь Маркова; он подробно рассмотрен в литературе¹⁾. Общий случай может быть описан следующим образом.

¹⁾ См. В. И. Романовский, «Дискретные цепи Маркова», Гостехиздат, 1949. (Прим. ред.)

Существует конечное число возможных «состояний» системы S_1, S_2, \dots, S_n . Кроме того, имеется совокупность вероятностей переходов $p_i(j)$, т. е. вероятностей того, что система, находящаяся в состоянии S_i , перейдет затем в состояние S_j . Чтобы представить при помощи этой цепи Маркова источник сообщений, достаточно только предположить, что при каждом переходе из одного состояния в другое создается одна буква. Состояния будут соответствовать «остатку влияния» предшествовавших букв.



Фиг. 3. График, соответствующий источнику в примере Б.



Фиг. 4. График, соответствующий источнику в примере В.

Все это может быть изображено графически, как показано на фиг. 3 и 4. Состояниями являются точки разветвления, а вероятности переходов и создаваемые при этом буквы указаны около соответствующих линий. Фиг. 3 относится к примеру Б раздела 2, фиг. 4 — к примеру В. На фиг. 3 имеется только одно состояние, так как последующие буквы независимы друг от друга. На фиг. 4 имеется столько же состояний, сколько букв. При учете трехбуквенных сочетаний было бы самое большее n^2 состояний, соответствующих возможным парам букв, предшествовавших выбираемой.

4. ЭРГОДИЧЕСКИЕ И СМЕШАННЫЕ ИСТОЧНИКИ

Как указано выше, дискретный источник может быть для наших целей представлен цепью Маркова. Среди возможных дискретных цепей Маркова имеется одна группа с особыми свойствами, имеющими значение в теории связи. Этот особый класс состоит из эргодических цепей; соответствующие источники также называются эргодическими. Хотя точное определение эргодического процесса несколько сложно, общая идея проста.

В случае эргодического процесса каждая создаваемая процессом последовательность имеет одни и те же статистические свойства. Так, частоты букв, частоты двухбуквенных сочетаний и т. д., полученные из частных последовательностей, будут стремиться по мере увеличения длины последовательностей к определенным пределам независимо от выбора частной последовательности. В действительности это верно не для всякой последовательности, но совокупность последовательностей, для которых это неверно, обладает

вероятностью 0. Грубо говоря, эргодичность означает статистическую однородность.

Свойство эргодичности связано со структурой соответствующего графика. Процесс будет эргодическим, если график обладает следующими двумя свойствами:

1. График не распадается на две изолированные части A и B , такие, что от одной точки разветвления в части A нельзя было бы перейти вдоль линии графика в направлении стрелок в точки части B , а также невозможно было бы и из точек части B попасть в точки части A .

2. Замкнутая последовательность линий графика, стрелки которых ориентированы в одном направлении, называется контуром. Под «длиной» контура понимается число линий, из которых он состоит. Второе свойство требует, чтобы наибольший общий делитель длин всех контуров графика равнялся единице.

Если первое условие удовлетворено, а второе нарушено тем, что общий делитель $d > 1$, то последовательности имеют некоторого рода периодическую структуру. Различные последовательности распадаются на d различных классов, которые в статистическом отношении одинаковы за исключением сдвига начала (т. е. выбора того, какую букву последовательности назвать первой). Путем смещения на величину от нуля до $d-1$ каждая последовательность может быть сделана статистически эквивалентной любой другой.

Простым примером при $d=2$ является следующее. Имеются три возможные буквы a, b, c . За буквой a следует либо b , либо c с вероятностями $1/3$ и $2/3$ соответственно. За b и за c всегда следует буква a . Тогда типичная последовательность имеет вид

$a b a c a c a c a b a c a b a b a c a c$

Такие случаи не имеют большого значения для нашей работы.

Если нарушено первое условие, то график может быть разделен на некоторое число частных графиков, каждый из которых удовлетворяет первому условию.

Будем предполагать, что второе условие также выполняется для всех частных графиков. В этом случае имеем то, что может быть названо «смешанным» источником, составленным из некоторого числа «чистых» составляющих. Составляющие соответствуют различным частным графикам. Если L_1, L_2, L_3, \dots — составляющие источники, то можно написать

$$L = p_1 L_1 + p_2 L_2 + p_3 L_3 + \dots$$

где p_i — вероятность составляющего источника L_i .

Физически дело обстоит таким образом. Имеется несколько различных источников L_1, L_2, L_3, \dots , каждый из которых имеет однородную статистическую структуру (т. е. является эргодическим). Мы не знаем априори, который будет использован, но раз последовательность началась с данной «чистой» составляющей L_i ,

она будет продолжаться бесконечно согласно своей статистической структуре.

Для примера можно взять два процесса, описанных выше, и предположить $p_1=0,2$ и $p_2=0,8$. Последовательность от смешанного источника

$$L = 0,2 L_1 + 0,8 L_2$$

может быть получена путем выбора первым L_1 или L_2 с вероятностями 0,2 и 0,8, а затем создания последовательности, определенной этим выбором.

Если не оговорено противное, то будем предполагать, что источник является эргодическим. Такое предположение позволяет отождествлять средние значения, взятые по последовательности, со средними значениями, взятыми по совокупности возможных последовательностей (вероятность отклонения равна нулю). Например, относительная частота буквы A в частной бесконечной последовательности будет с вероятностью единица равняться ее относительной частоте в совокупности последовательностей.

Если P_i — вероятность состояния i , а $p_i(j)$ — вероятность перехода в состояние j , то, чтобы процесс был стационарным, P_i должно, очевидно, удовлетворять условиям равновесия

$$P_j = \sum_i P_i p_i(j).$$

В эргодическом случае можно показать, что при любых начальных условиях вероятности $P_j(N)$ пребывания в состоянии j после N символов приближаются к равновесным значениям при $N \rightarrow \infty$.

5. ВЫБОР, НЕОПРЕДЕЛЕННОСТЬ И «ЭНТРОПИЯ»

Дискретный источник сообщений был представлен выше в виде цепи Маркова. Возникает вопрос, можно ли определить величины, которые измеряли бы в определенном смысле, сколько ^{данных} «создаст» такой процесс, или лучше, с какой скоростью ^{создаются} данные? *как?*

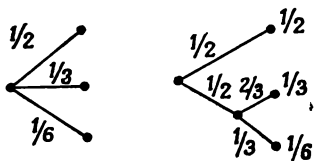
Пусть имеется набор возможных событий, вероятности появления которых суть p_1, p_2, \dots, p_n . Эти вероятности известны, но это все, что знаем относительно того, какое произойдет событие. Можно ли найти меру того, чему равна «возможность выбора» или какова неопределенность исхода при выборе события из этой группы?

Если такая мера существует, — обозначим ее $H(p_1, p_2, \dots, p_n)$, — то целесообразно потребовать, чтобы она обладала следующими свойствами:

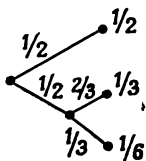
1. H должна быть непрерывна относительно p_i .
2. Если все p_i одинаковы, $p_i = \frac{1}{n}$, то H должна быть монотонно возрастающей функцией от n . В случае равновероятных событий

имеет место бóльшая возможность выбора или бóльшая неопределенность, чем в тех случаях, когда имеются и более вероятные и менее вероятные события.

3. Если выбор распадается на два последовательных этапа, то исходное H будет взвешенной суммой индивидуальных значений H . Смысл этого иллюстрирует фиг. 5. Слева имеются три возможности



Фиг. 5. Различные случаи выбора.



с вероятностями $p_1 = \frac{1}{2}$, $p_2 = \frac{1}{3}$, $p_3 = \frac{1}{6}$.

Справа сначала выбираем между двумя возможностями с вероятностями $\frac{1}{2}$ и в случае второй возможности делаем еще выбор между возможностями с вероятностями $\frac{2}{3}$ и $\frac{1}{3}$. Окончательные результаты имеют те же самые вероятности, как и прежде.

В этом частном случае требуется, чтобы

$$H\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{6}\right) = H\left(\frac{1}{2}, \frac{1}{2}\right) + \frac{1}{2} H\left(\frac{2}{3}, \frac{1}{3}\right).$$

Коэффициент $\frac{1}{2}$ появился потому, что выбор на втором этапе происходит только в половине общего числа случаев.

В Приложении 2 обосновывается следующий вывод:

Т е о р е м а 2

Единственной формой H , удовлетворяющей трем указанным предположениям, является форма

$$H = -K \sum_{i=1}^n p_i \log p_i,$$

где K — положительная постоянная.

Эта теорема, как и предположения, используемые для ее доказательства, имеют для данной теории второстепенное значение. Она дана главным образом для того, чтобы подкрепить закономерность наших дальнейших определений.

Величины вида $H = -\sum p_i \log p_i$ (постоянная K определяет только единицу измерения) играют центральную роль в теории передачи сообщений в качестве мер возможности выбора и неопределенности. Форма величины H такова же, как и энтропии, определяемой в статистической механике, где p_i — вероятность того, что система находится в ячейке i своего фазового пространства.

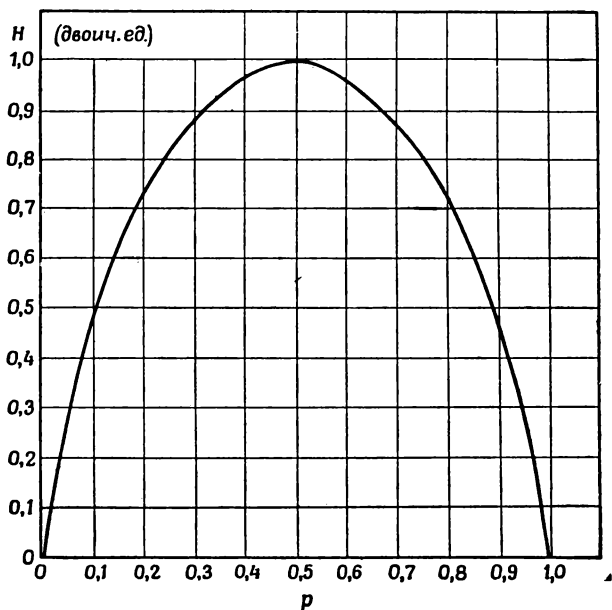
Условимся называть $H = -\sum p_i \log p_i$ «энтропией»¹⁾ совокупности вероятностей p_1, \dots, p_n . Если случайная переменная обозначена

¹⁾ Автор вводит термин «энтропия» на основании чисто внешнего сходства выражения для введенной им величины H с выражением для энтропии в общепринятом значении. Поскольку с понятием энтропии в статистической

x , то будем обозначать «энтропию» через $H(x)$; таким образом, x — не аргумент функции, а лишь знак, отличающий ее от $H(y)$, т. е. от «энтропии» случайной величины y .

На фиг. 6 представлена в виде функции от p «энтропия» в случае двух возможностей с вероятностями p и $q=1-p$, а именно:

$$H = -(p \log p + q \log q).$$



Фиг. 6. «Энтропия» в случае двух возможностей с вероятностями p и $(1-p)$.

Величина H обладает рядом интересных свойств, которые подтверждают правильность применения ее в качестве рациональной меры возможности выбора.

1. $H=0$ только в том случае, если все вероятности p_i , кроме одной, равны нулю, а эта единственная вероятность равна единице. Таким образом, H равна нулю только в случае полной определенности исхода опыта. В остальных случаях H больше нуля.

физике связано вполне определенное физическое содержание, то во избежание возможных недоразумений из-за формалистической терминологии автора в дальнейшем слово «энтропия» поставлено в кавычки. Следует иметь в виду, что в данном случае «энтропия» есть не больше, чем краткое название величины $H = \sum p_i \log p_i$, где p_i — вероятность появления некоторого события i . (Прим. ред.)

2. При заданном n H максимальна и равна $\log n$, когда все p_i одинаковы $\left(\frac{1}{n}\right)$. Это и интуитивно соответствует состоянию наибольшей неопределенности.

3. Пусть имеются два события x и y с m возможностями для первого и n — для второго. Пусть $p(i, j)$ означает вероятность совместного появления i -го значения для первого и j -го значения для второго события. «Энтропия» совместного события равна

$$H(x, y) = - \sum_{i, j} p(i, j) \log p(i, j),$$

тогда как

$$H(x) = - \sum_{i, j} p(i, j) \log \sum_j p(i, j),$$

$$H(y) = - \sum_{i, j} p(i, j) \log \sum_i p(i, j).$$

Легко показать, что

$$H(x, y) \leq H(x) + H(y),$$

причем равенство имеет место только в случае независимых событий [т. е. если $p(i, j) = p(i)p(j)$].

4. Всякое изменение в сторону выравнивания вероятностей p_1, p_2, \dots, p_n увеличивает H . Так, если $p_1 < p_2$ и увеличиваем p_1 , одновременно уменьшая p_2 на такую же величину, так что p_1 и p_2 приближаются друг к другу, то H увеличивается.

В более общем виде, если над вероятностями осуществляется операция «выравнивания» вида

$$p'_i = \sum_j a_{ij} p_j,$$

где $\sum_i a_{ij} = \sum_j a_{ij} = 1$ и все $a_{ij} \geq 0$, то H увеличивается. (За исключением того частного случая, в котором такое преобразование сводится к одной только перестановке p_j , когда H , конечно, сохраняется неизменной.)

5. Пусть имеются два случайных события x и y , как в пункте 3, но обязательно независимых. Для каждого частного значения i , которое может принимать x , имеется условная вероятность $p_i(j)$ того, что y имеет значение j . Она равна

$$p_i(j) = \frac{p(i, j)}{\sum_j p(i, j)}.$$

Мы определяем *условную энтропию* величины y , $H_x(y)$ как среднее значение «энтропии» y для каждого значения x , вычисленное с учетом весов, соответствующих вероятностям частных значений x .

Она равна

$$H_x(y) = - \sum_{i,j} p(i, j) \log p_i(j).$$

Эта величина показывает, какова в среднем неопределенность значения y при известном x . Подставляя значение $p_i(j)$, получим

$$H_x(y) = - \sum_{i,j} p(i, j) \log p(i, j) + \sum_{i,j} p(i, j) \log \sum_j p(i, j) = H(x, y) - H(x)$$

или

$$H(x, y) = H(x) + H_x(y).$$

Неопределенность (или «энтропия») совместного события (x, y) есть неопределенность события x плюс неопределенность события y , когда x известно.

6. Из 3 и 5-го пунктов имеем

$$H(x) + H(y) \geq H(x, y) = H(x) + H_x(y).$$

Отсюда

$$H(y) \geq H_x(y).$$

Неопределенность события y никогда не возрастает вследствие знания события x . Она уменьшается, если только события x и y не являются независимыми. В противном случае она не изменяется.

6. «ЭНТРОПИЯ» ИСТОЧНИКА СООБЩЕНИЙ

Рассмотрим дискретный источник с конечным числом состояний, вроде рассмотренных выше. Для каждого возможного состояния i имеется совокупность вероятностей $p_i(j)$ создания различных возможных символов j . Для каждого состояния существует «энтропия» H_i . «Энтропия» источника определяется как среднее значение этих H_i , каждому из которых приписан вес, в соответствии с вероятностью появления соответствующего состояния

$$H = \sum_i P_i H_i = - \sum_{i,j} P_i p_i(j) \log p_i(j).$$

Это — «энтропия» источника на символ текста. Если процесс Маркова развивается с определенной скоростью, то можно говорить также об «энтропии» в секунду

$$H' = \sum_i f_i H_i,$$

где f_i — средняя частота (появлений в секунду) состояния i . Очевидно,

$$H' = mH,$$

где m — среднее число символов, создаваемых за 1 сек.

H или H' измеряют количество данных, создаваемое источником на символ или за секунду. Если в качестве основания логарифмов

выбрано 2, то они будут выражены в двоичных единицах на символ или за секунду.

Если последовательные символы независимы, то H просто равняется — $\sum p_i \lg p_i$, где p_i — вероятность символа i . Предположим, что в этом случае мы рассматриваем длинное сообщение из N символов. Оно будет содержать с большой вероятностью около $p_1 N$ появлений первого символа, $p_2 N$ появлений второго и т. д. Отсюда вероятность данного частного сообщения будет приближенно равна

$$p = p_1^{p_1 N} p_2^{p_2 N} \dots p_n^{p_n N}$$

или

$$\log p = N \sum_i p_i \log p_i,$$

$$\log p = -NH,$$

$$H = \frac{\log \frac{1}{p}}{N}.$$

Поэтому H приближенно равна логарифму обратной величины вероятности типичной длинной последовательности, деленному на число символов в последовательности. Тот же результат сохраняется и для любого источника.

Формулируя более точно, имеем (см. Приложение 3):

Т е о р е м а 3

Для любых заданных $\epsilon > 0$ и $\delta > 0$ можно найти такое N_0 , что последовательности любой длины $N \geq N_0$ распадаются на два класса:

1. Группа последовательностей, общая вероятность которых меньше, чем ϵ .

2. Остаток, все члены которого обладают вероятностями, удовлетворяющими неравенству

$$\left| \frac{\log \frac{1}{p}}{N} - H \right| < \delta.$$

Другими словами, почти достоверно, что $\frac{\log \frac{1}{p}}{N}$ весьма близко к H , когда N велико.

Близкий результат справедлив для любого числа последовательностей с различными вероятностями.

Рассмотрим опять последовательности длины N . Расположим их в порядке уменьшения вероятностей. Введем $n(q)$ — число последовательностей, которые мы должны взять из этой совокупности, начиная с наиболее вероятной последовательности, чтобы

накопить полную вероятность q для взятых нами последовательностей.

Теорема 4

$$\lim_{N \rightarrow \infty} \frac{\log n(q)}{N} = H,$$

когда q не равно нулю или единице.

$\log n(q)$ можно истолковать как число двоичных единиц, требуемых для описания последовательности, когда рассматриваются только наиболее вероятные последовательности с общей вероятностью q . Тогда $\frac{\log n(q)}{N}$ есть число двоичных единиц на символ, необходимых для описания последовательностей. Теорема гласит, что для больших N оно не зависит от q и равно H . Быстрота возрастания логарифма числа сравнительно вероятных последовательностей определяется величиной H , независимо от истолкования термина «сравнительно вероятный». Благодаря этим результатам, доказанным в Приложении 3, можно в большинстве случаев рассматривать длинные последовательности, как если бы их было 2^{HN} , каждая с вероятностью 2^{-HN} .

Следующие две теоремы показывают, что H и H' могут быть определены предельными переходами непосредственно из статистики последовательностей сообщений без рассмотрения вероятностей состояний и вероятностей переходов между состояниями.

Теорема 5

Пусть $p(B_i)$ — вероятность появления на выходе источника последовательности символов B_i . Пусть

$$G_N = -\frac{1}{N} \sum_i p(B_i) \log p(B_i),$$

где суммирование распространяется на все последовательности B_i , содержащие N символов. Тогда G_N является монотонно убывающей функцией от N и

$$\lim_{N \rightarrow \infty} G_N = H.$$

Теорема 6

Пусть $p(B_i, S_j)$ — вероятность появления последовательности B_i , сопровождаемой символом S_j , а $p_{B_i}(S_j) = \frac{p(B_i, S_j)}{p(B_i)}$ — условная вероятность того, что S_j следует за B_i . Пусть

$$F_N = -\sum_{i,j} p(B_i, S_j) \log p_{B_i}(S_j),$$

где суммирование распространено по всем группам B_i из $N-1$ символов и по всем символам S_j . Тогда F_N является монотонно убывающей функцией от N

$$F_N = NG_N - (N-1)G_{N-1}, \quad G_N = \frac{1}{N} \sum_1^N F_N, \quad F_N \leq G_N$$

и

$$\lim_{N \rightarrow \infty} F_N = H.$$

Все эти результаты получены в Приложении 3. Они показывают, что ряд приближений к H может быть найден путем рассмотрения одной только статистической структуры последовательностей, охватывающих $1, 2, \dots, N$ символов. F_N является наилучшим приближением. В самом деле, F_N есть «энтропия» N -го приближения к источнику рассмотренного выше типа. Если статистическое влияние, распространяющееся больше чем на N символов, отсутствует, т. е. если условная вероятность появления следующего символа, при условии знания предшествовавших $(N-1)$ символов, не изменяется при ознакомлении с любыми символами, стоящими ранее, то $F_N = H$.

F_N является, конечно, условной «энтропией» следующего символа, когда известны предыдущие $(N-1)$ символов, тогда как G_N — «энтропия» на символ для групп из N символов.

Отношение «энтропии» источника к ее максимальному значению, которое она может иметь при тех же символах, называется *относительной энтропией* источника. Как будет показано ниже, это — максимальная сжатость, которая может быть достигнута кодированием при помощи того же самого алфавита.

Единица минус относительная «энтропия» есть *избыточность*. Избыточность обычного английского текста, если не рассматривать статистическую структуру, относящуюся более чем к 8 буквам, составляет примерно 50%. Это значит, что когда пишут по-английски, то половина знаков текста определяется структурой языка и лишь половина выбирается по желанию пишущего.

Число 50% было найдено несколькими независимыми методами, которые все дают сходные результаты. Одним из таких методов было вычисление «энтропии» приближений к английскому тексту. Другой состоял в исключении из образцов английского текста некоторой части букв, после чего делалась попытка их восстановить. Если бы их удалось восстановить, когда 50% текста исключено, избыточность должна бы быть больше, чем 50%. Третий метод связан с некоторыми известными выводами криптографии.

7. ПРЕДСТАВЛЕНИЕ ПРОЦЕССОВ КОДИРОВАНИЯ И ДЕКОДИРОВАНИЯ

Теперь нужно представить математически операции, выполняемые передатчиком и приемником с целью кодирования и декодирования сообщений. Каждый из них будет называться дискретным четырехполюсником. На вход четырехполюсника поступает последовательность входных символов, а на выходе получается последовательность выходных символов. В общем случае выходной эффект четырехполюсника зависит не только от наличного входного символа, но также и от предыдущих. Предположим, что существует конечное число m возможных состояний четырехполюсника и что его выходной эффект является функцией наличного состояния и наличного входного символа. Следующее состояние будет функцией этих двух величин.

Поэтому четырехполюсник может быть описан двумя функциями:

$$y_n = f(x_n, \alpha_n), \quad \alpha_{n+1} = g(x_n, \alpha_n),$$

где x_n есть n -й входной символ, α_n — состояние четырехполюсника в момент введения n -го входного символа, y_n — выходной символ (или последовательность выходных символов), создаваемый, когда на входе действует символ x_n и четырехполюсник находится в состоянии α_n .

Если выходные символы четырехполюсника можно отождествить со входными символами другого четырехполюсника, то четырехполюсники могут быть соединены последовательно, в результате чего получится новый четырехполюсник.

Если существует второй четырехполюсник, который работает от выхода первого четырехполюсника и восстанавливает исходный входной сигнал, то первый четырехполюсник называется несингулярным, а второй — ему обратным.

Теорема 7

Выход четырехполюсника с конечным числом состояний, возбуждаемого статистическим источником также с конечным числом состояний, является статистическим источником с конечным числом состояний и с «энтропией» (на единицу времени), меньшей или равной «энтропии» на входе. Если четырехполюсник не сингулярен, «энтропии» равны.

Пусть α представляет собой состояние источника, который дает последовательность символов x_i , и пусть β — состояние четырехполюсника, который создает на выходе группы символов y_i . Комбинированная система может быть представлена «пространством результирующих состояний» пар (α, β) . Две точки в этом пространстве (α_1, β_1) и (α_2, β_2) соединяются линией, если состоя-

ние α_1 может создать x , который превращает β_1 в β_2 , и эта линия дает вероятность этого x в данном случае. Линия обозначается группой из y_1 символов, создаваемых четырехполюсником.

«Энтропия» выходного эффекта может быть вычислена как взвешенная сумма по всем состояниям. Если суммировать сначала по β , то получающиеся члены меньше или равны соответствующим членам для α , следовательно, «энтропия» не возрастает. Присоединим выход несингулярного четырехполюсника к обратному четырехполюснику. Если H'_1 , H'_2 и H'_3 суть «энтропии» выходных эффектов соответственно источника, первого и второго четырехполюсников, то $H'_1 \geq H'_2 \geq H'_3 = H'_1$ и, следовательно:

$$H'_1 = H'_2.$$

Пусть имеется система с ограничениями, наложенными на возможные последовательности того типа, который можно представить линейным графиком фиг. 2. Если вероятности $p_{ij}^{(s)}$ приписаны различным линиям, соединяющим состояние i с состоянием j , то эта система будет источником. Существует один частный способ назначить вероятности, который дает максимум «энтропии» (см. Приложение 4).

Теорема 8

Пусть система с ограничениями, рассматриваемая как канал, обладает пропускной способностью $C = \log W$. Если положим

$$p_{ij}^{(s)} = \frac{B_j}{B_i} W^{-l_{ij}^{(s)}},$$

где $l_{ij}^{(s)}$ — длительность s -го символа, ведущего от состояния i к состоянию j , а B_i удовлетворяет условию

$$B_i = \sum_{s,j} B_j W^{-l_{ij}^{(s)}},$$

то «энтропия» H имеет максимум и равна пропускной способности C .

Путем надлежащего назначения вероятностей переходов «энтропия» символов в канале может быть доведена до максимума, равного пропускной способности канала.

8. ОСНОВНАЯ ТЕОРЕМА ДЛЯ КАНАЛА БЕЗ ШУМОВ

Проверим теперь правильность интерпретации величины H , как скорости создания данных, путем доказательства того, что H определяет пропускную способность канала, необходимую при наиболее эффективном кодировании.

Т е о р е м а 9

Пусть источник имеет «энтропию» H (двоичных единиц на символ), а канал обладает пропускной способностью C (двоичных единиц в 1 сек.). Тогда можно закодировать сообщения на выходе источника таким образом, чтобы передавать символы по каналу со средней скоростью $\frac{C}{H} - \epsilon$ символов в 1 сек., где ϵ — сколько угодно мало. Передавать со средней скоростью, большей чем $\frac{C}{H}$, невозможно.

Обратная часть теоремы, утверждающая, что нельзя превзойти скорости $\frac{C}{H}$, может быть доказана, если заметить, что «энтропия» на входе канала за 1 сек. равна «энтропии» источника, так как передатчик должен быть несингулярным четырех-полюсником, и что эта «энтропия» не может превзойти пропускной способности канала. Отсюда $H' \leq C$ и число символов в 1 сек. равно $\frac{H'}{H} \leq \frac{C}{H}$.

Первая часть теоремы будет доказана двумя различными способами. Первый способ состоит в рассмотрении совокупности всех последовательностей из N символов, создаваемых источником. При большом N можно разделить их на две группы, одна из которых содержит меньше чем $2^{(H+\eta)N}$ членов, а вторая меньше чем 2^{RN} членов (где R — логарифм числа различных символов) и имеет полную вероятность, меньшую μ . Если N возрастает, η и μ приближаются к нулю. Число сигналов в канале, каждый длительностью T , больше чем $2^{(C-\theta)T}$, причем θ мало, когда T велико. Если выбрать

$$T = \left(\frac{H}{C} + \lambda \right) N,$$

то найдется достаточное число последовательностей канальных символов для группы, обладающей высокой вероятностью, когда N и T достаточно велики (как бы ни было мало λ), а также несколько добавочных последовательностей символов. Группа последовательностей с высокой вероятностью произвольным, взаимно однозначным образом, кодируется в эту совокупность. Остающиеся последовательности представляются более длинными, начинающимися и заканчивающимися одной из последовательностей, не использованных для группы с высокой вероятностью. Эта особая последовательность играет роль стартстопного сигнала другой кодовой комбинации.

Между ними сохраняется временной интервал, необходимый для образования достаточного числа различных последователь-

ностей для всех маловероятных сообщений. Для этого требуется

$$T_1 = \left(\frac{R}{C} + \varphi \right) N,$$

где φ мало.

Средняя скорость передачи символов сообщения в 1 сек. будет тогда больше, чем

$$\left[(1 - \delta) \frac{T}{N} + \delta \frac{T_1}{N} \right]^{-1} = \left[(1 - \delta) \left(\frac{H}{C} + \lambda \right) + \delta \left(\frac{R}{C} + \varphi \right) \right]^{-1}$$

При возрастании $N\delta$, λ и φ стремятся к нулю, а скорость приближается к $\frac{C}{H}$.

Другой подход к такому кодированию и иной метод доказательства теоремы состоят в следующем. Расположим сообщения длиной N в порядке убывания вероятностей, и пусть эти вероятности будут $p_1 \geq p_2 \geq p_3 \dots \geq p_n$.

Пусть $P_s = \sum_1^{s-1} p_i$; это значит, что P_s есть накопленная вероятность вплоть до p_{s-1} включительно. Произведем сначала кодирование по двоичной системе. Кодовая комбинация для сообщения s получается путем разложения P_s как двоичного числа. Разложение будет содержать m_s позиций, где m_s есть целое число, удовлетворяющее соотношению

$$\lg_2 \frac{1}{p_s} \leq m_s < 1 + \lg_2 \frac{1}{p_s}.$$

Таким образом, высоковероятные сообщения представляются короткими кодовыми комбинациями, а маловероятные — длинными. Из этих неравенств вытекает

$$\frac{1}{2^{m_s}} \leq p_s < \frac{1}{2^{m_s-1}}.$$

Новая комбинация для P_s будет отличаться от всех последующих одним или более из своих m_s знаков, так как все остающиеся P_i , по крайней мере, на $\frac{1}{2^{m_s}}$ больше, и потому их двоичное разложение отличается первыми m_s знаками. Следовательно, все кодовые комбинации различны, и по ним можно восстановить сообщение. Если каналные последовательности не всегда являются последовательностями двоичных знаков, они могут быть записаны двоичными знаками произвольным образом, а двоичный код преобразован в сигналы, годные для канала.

Среднее число H_1 двоичных знаков, употребляемых на символ первоначального сообщения, легко определить. Мы имеем

$$H_1 = \frac{1}{N} \sum m_s p_s.$$

Но

$$\frac{1}{N} \sum \left(\log_2 \frac{1}{p_s} \right) p_s \leq \frac{1}{N} \sum m_s p_s < \frac{1}{N} \sum \left(1 + \log_2 \frac{1}{p_s} \right) p_s,$$

и поэтому

$$G_N \leq H_1 < G_N + \frac{1}{N}.$$

Когда N увеличивается, G_N сходится к H — «энтропии» источника, а H_1 сходится к H .

Отсюда видно, что неэффективность кодирования в случае конечного времени запаздывания N символов не должна быть больше, чем $\frac{1}{N}$ плюс разность между истинной «энтропией» H и «энтропией» G_N , вычисленной для последовательностей длины N . Избыточное относительное время, потребное сверх идеального случая, будет поэтому меньше, чем

$$\frac{G_N}{H} + \frac{1}{HN} - 1.$$

Этот метод кодирования в сущности совпадает с методом, независимо найденным Р. М. Фэнно. Его метод состоит в расположении сообщений длины N в порядке убывающих вероятностей. Этот ряд делится на две группы, по возможности с равными вероятностями. Если сообщение относится к первой группе, его первая двоичная цифра будет 0, в противном случае — 1. Группа подобным же образом разделяется на подгруппы примерно равной вероятности, и частная подгруппа определяет второй двоичный знак. Этот процесс продолжается до тех пор, пока не получатся подгруппы, содержащие только по одному сообщению. Легко видеть, что за исключением незначительных отличий (в общем случае в последней цифре) это приводит к тому же самому положению вещей, как и описанный выше арифметический процесс.

9. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ И ПРИМЕРЫ

Для передачи максимальной мощности от генератора в нагрузку в общем случае применяется трансформатор, который делает сопротивление генератора, наблюдаемое со стороны нагрузки, равным сопротивлению нагрузки. В рассматриваемом случае положение вещей, грубо говоря, аналогично. Четырехполюсник, осу-

шествующий кодирование, согласовывает источник с каналом в статистическом смысле. Источник, рассматриваемый через четырехполосник со стороны канала, должен иметь ту же самую статистическую структуру, какую имеет источник, обеспечивающий максимум «энтропии» в канале.

Содержание теоремы 9 сводится к тому, что, хотя точное согласование в общем случае невозможно, к нему можно подойти сколь угодно близко. Отношение действительной скорости передачи к пропускной способности канала может быть названо эффективностью кодирующей системы. Она, конечно, равна отношению действительной «энтропии» канальных символов к максимально возможной «энтропии».

Вообще говоря, идеальное или близкое к этому кодирование требует длительных временных задержек в передатчике и приемнике. В случае отсутствия шумов, который и рассматривается, главное назначение этих задержек состоит в рациональном согласовании вероятностей с соответствующими длительностями последовательностей. При хорошем коде логарифм обратной величины вероятности длинного сообщения должен быть пропорционален длительности соответствующего сигнала; действительно

$$\left| \log \frac{1}{p} - C \right|$$

должно быть мало для всех длинных сообщений за исключением их небольшой части.

Если источник может давать только одно определенное сообщение, его «энтропия» равна нулю, и канал не нужен. Например, счетная машина, спроектированная для вычисления последовательных цифр числа π , дает определенную последовательность без всяких элементов случайности. Для «передачи» этой последовательности в другую точку не нужно никакого канала. Во второй точке можно построить другую машину, вычисляющую ту же самую последовательность. Однако это может быть непрактично. В этом случае можно игнорировать статистические сведения об источнике или только часть их. Можно рассматривать цифры числа π как случайную последовательность и сконструировать систему, способную передавать любую последовательность цифр.

Подобным же образом можно использовать некоторые из статистических сведений об английском тексте при составлении кода. В этом случае рассматривается источник с максимальной «энтропией», подчиненный статистическим условиям, которые пожелали сохранить. «Энтропия» этого источника определяет необходимую и достаточную пропускную способность канала. В примере с числом π оставлены только те сведения, что все цифры выбираются из ряда 0,1,...,9. В случае английского текста можно пожелать использо-

вать статистическую экономию, возможную благодаря знанию частот букв, но ничего больше. Источник с максимальной «энтропией» будет тогда первым приближением к английскому тексту и его «энтропия» определит необходимую пропускную способность канала.

В качестве простого примера использования некоторых из полученных результатов рассмотрим источник, создающий последовательность букв, выбранных из ряда A, B, C, D с вероятностями $1/2, 1/4, 1/8, 1/8$, причем последовательные символы выбираются независимо. Имеем

$$H = - \left(\frac{1}{2} \log_2 \frac{1}{2} + \frac{1}{4} \log_2 \frac{1}{4} + \frac{2}{8} \log_2 \frac{1}{8} \right) = \frac{7}{4} \frac{\text{двоичных единиц}}{\text{символ}}$$

Таким образом, для кодирования сообщений этого источника двоичными знаками в пределе достаточно в среднем $7/4$ знака на символ.

В этом случае можно действительно достигнуть предельного значения, применяя следующий код (полученный по методу второго доказательства теоремы 9):

A	0
B	10
C	110
D	111

Среднее число двоичных знаков, применяемых для кодирования последовательности из N символов, будет

$$N \left(\frac{1}{2} \times 1 + \frac{1}{4} \times 2 + \frac{2}{8} \times 3 \right) = \frac{7}{4} N.$$

Легко видеть, что двоичные знаки 0,1 имеют вероятности $1/2, 1/2$, так что «энтропия» для кодированных последовательностей равна одной двоичной единице на символ. Так как в среднем имеем $7/4$ двоичных знаков на букву оригинала, то «энтропия» на единицу времени будет той же самой. Максимально возможная «энтропия» для первоначального ряда равна $\log_2 4 = 2$ и имеет место, когда A, B, C, D обладают вероятностями $1/4, 1/4, 1/4, 1/4$. Отсюда относительная «энтропия» равна $7/8$. Мы можем перевести двоичные последовательности в первоначальный ряд символов в соотношении 2 к 1 по следующей таблице:

00	A'
01	B'
10	C'
11	D'

Этот двойной процесс кодирует первоначальное сообщение в те же самые символы, но со средним коэффициентом сжатия $7/8$.

В качестве второго примера рассмотрим источник, дающий последовательность букв A и B с вероятностями p для A и q для B . Если $p \ll q$, имеем

$$\begin{aligned} H &= -\log p^p (1-p)^{1-p} = \\ &= -p \log p (1-p)^{\frac{1-p}{p}} = \\ &= p \log \frac{e}{p}. \end{aligned}$$

В этом случае можно построить хорошую систему кодирования сообщений в канале, передающем только 0,1, путем посылки специальной последовательности, скажем 0000, для редкого символа A и затем последовательности, указывающей число букв B , следующих за ним. Это число может быть указано путем представления в двоичной системе, причем все числа, содержащие специальную последовательность, исключаются. Все числа до 16 изображаются как обычно; 16 передается следующим после шестнадцати двоичным числом, которое не содержит четырех нулей, а именно: $17 = 10001$ и т. д.

Можно показать, что при $p \rightarrow 0$ кодирование приближается к идеальному, если только длина специальной последовательности выбрана правильно.

Глава II

ДИСКРЕТНЫЙ КАНАЛ С ШУМАМИ

10. ПРЕДСТАВЛЕНИЕ ДИСКРЕТНОГО КАНАЛА С ШУМАМИ

Рассмотрим теперь случай, когда на сигнал в процессе передачи или на концах тракта воздействуют шумы. Это означает, что принимаемый сигнал не обязательно тот же самый, что посланный передатчиком. Можно различать два случая. Если данный переданный сигнал всегда создает тот же самый принятый сигнал, т. е. принятый сигнал является определенной функцией переданного сигнала, то такой эффект может быть назван искажением. Если эта функция имеет обратную — никакие два переданных сигнала не создают одинаковых принятых сигналов, — то искажения могут быть скорректированы, по крайней мере принципиально, просто путем выполнения обратного функционального преобразования принятого сигнала.

Интересен случай, когда сигнал испытывает не всегда одинаковое изменение при передаче. Тогда можно считать принятый сигнал E функцией переданного сигнала S и другой переменной — шумов N

$$E = f(S, N).$$

Шумы рассматриваются как случайная переменная, точно так же как раньше рассматривалось сообщение. В общем случае шумы могут быть представлены подходящим стохастическим процессом. Наиболее общий тип дискретного канала с шумами, какой будет рассмотрен, является обобщением ранее описанного свободного от шумов канала с конечным числом состояний. Предположим, что число состояний конечно и имеется совокупность вероятностей

$$p_{\alpha, i}(\beta, j)$$

того, что если канал находится в состоянии α и передается символ i , то будет принят символ j , а канал перейдет в состояние β . Таким образом, α и β охватывают все возможные состояния, i — все возможные передаваемые сигналы, а j — все возможные принимаемые сигналы. Если последовательные символы подвергаются воздействию шумов независимо, имеется только одно состояние, и канал описывается совокупностью вероятностей переходов $p_i(j)$ (вероятность того, что переданный сигнал i будет принят как j).

Когда на входе канала с шумами действует некоторый процесс сообщений, следует рассматривать два статистических процесса,

описывающих источник и шумы. Соответственно этому может быть вычислен ряд различных «энтропий».

Во-первых, существует «энтропия» $H(x)$ источника или «энтропия» на входе канала (они равны, если передатчик не сингулярен). «Энтропия» на выходе канала, т. е. «энтропия» принимаемого сигнала, будет обозначаться через $H(y)$. В случае отсутствия шумов $H(y) = H(x)$. «Энтропия» совместного события для входа и выхода будет обозначаться через $H(x, y)$. Наконец, имеются две условные «энтропии» $H_x(y)$ и $H_y(x)$, «энтропия» на выходе при известном эффекте на входе, и наоборот.

Между этими величинами имеет место соотношение

$$H(x, y) = H(x) + H_x(y) = H(y) + H_y(x).$$

Все эти «энтропии» могут измеряться за 1 сек. или на символ.

11. НЕНАДЕЖНОСТЬ И ПРОПУСКНАЯ СПОСОБНОСТЬ КАНАЛА

При наличии шумов, вообще говоря, невозможно на основании принятого сигнала восстановить исходное сообщение или переданный сигнал *с полной определенностью*. Однако имеются некоторые способы передачи сообщений, которые являются оптимальными в отношении борьбы с шумами. Эта задача и рассматривается ниже.

Предположим, что имеется два возможных сигнала 0 и 1 и что передача осуществляется со скоростью в 1000 символов в 1 сек. с вероятностями $p_0 = p_1 = \frac{1}{2}$. Таким образом, данный источник создает сообщения со скоростью 1000 двоичных единиц в 1 сек. Во время передачи шумы вносят ошибки таким образом, что в среднем один из ста принятых сигналов неправилен (0 вместо 1 или 1 вместо 0).

Какова скорость передачи сообщений? Конечно, меньше, чем 1000 двоичных единиц в 1 сек., так как около одного процента принятых символов неправильны. Сразу же хочется сказать, что эта скорость составляет 990 двоичных единиц в 1 сек., т. е. просто вычесть число ошибок. Однако это неправильно, здесь не учитывается, что получатель не знает, где именно произошла ошибка.

Можно рассмотреть крайний случай и предположить, что шумы столь велики, что принятые символы совершенно не зависят от переданных. Вероятность приема 1 равна $\frac{1}{2}$ при передаче любого сообщения и аналогично для 0. Тогда около половины всех принятых символов будут правильными благодаря одной только случайности, и можно считать, что система способна передавать 500 двоичных единиц, в то время как в действительности данные вовсе не передаются. Можно было бы получить столь же «хорошую» передачу, отказавшись вообще от канала и подбрасывая монету в точке приема.

Очевидно, правильная поправка к количеству переданных данных равна количеству данных, отсутствующих в принятом сигнале, или иначе — неопределенности при приеме сигнала относительно того, что именно было в действительности передано. На основании предыдущих рассуждений относительно «энтропии» как меры неопределенности представляется рациональным использовать условную «энтропию» сообщения (при условии знания принятого сигнала) в качестве меры этих отсутствующих данных. Как увидим ниже, такое определение будет правильным.

Следуя этой идее, действительная скорость передачи может быть найдена путем вычитания из скорости создания сообщений (т. е. «энтропии» источника) условной «энтропии» $H_y(x)$

$$R = H(x) - H_y(x).$$

Условная «энтропия» $H_y(x)$ будет называться *ненадежностью*. Она измеряет среднюю неопределенность принятого сигнала.

В рассмотренном выше примере, если был принят нуль, то апостериорная вероятность того, что был передан нуль, равна 0,99, а что была передана единица, — 0,01. Если же была принята единица, то эти цифры поменяются местами.

Отсюда

$$H_y(x) = -(0,99 \lg 0,99 + 0,01 \lg 0,01) = 0,081 \frac{\text{двоичных единиц}}{\text{символ}}$$

или 81 двоичная единица в 1 сек. Можно сказать, что система передает со скоростью $1000 - 81 = 919$ двоичных единиц в 1 сек. В том крайнем случае, когда при передаче какого-то знака равновероятен прием как 0, так и 1, апостериорные вероятности равны $1/2$, $1/2$ и

$$H_y(x) = -\left(\frac{1}{2} \lg \frac{1}{2} + \frac{1}{2} \lg \frac{1}{2}\right) = 1 \frac{\text{двоичных единиц}}{\text{символ}}$$

или 1000 двоичных единиц в 1 сек. Скорость передачи, как и следовало ожидать, равна в этом случае нулю.

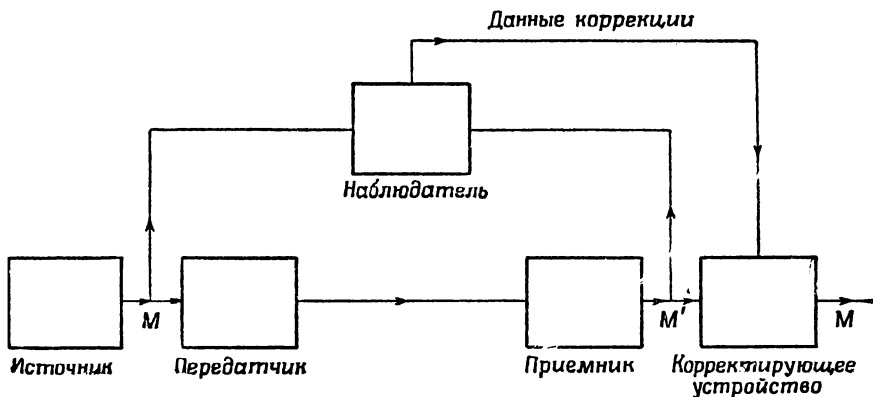
Следующая теорема дает непосредственную интуитивную интерпретацию ненадежности и подтверждает, что она является единственной подходящей мерой.

Рассмотрим систему связи и наблюдателя (или вспомогательный прибор), который может наблюдать как то, что передается, так и то, что принимается (с ошибками из-за шумов). Этот наблюдатель отмечает ошибки в воспроизводимом сообщении и передает данные в точку приема через «канал коррекции», чтобы дать возможность в точке приема исправить ошибки. Схематически это показано на фиг. 7

Теорема 10

Если канал коррекции обладает пропускной способностью, равной $H_y(x)$, то можно таким образом закодировать данные коррекции, чтобы передать их по этому каналу, и скорректировать все ошибки за исключением произвольно малой доли ϵ . Это невозможно, если пропускная способность канала коррекции меньше, чем $H_y(x)$.

Грубо говоря, $H_y(x)$ есть количество дополнительных данных, которые должны быть переданы за 1 сек. в точку приема для корректирования принятого сообщения.



Фиг. 7. Схема системы коррекции.

Для доказательства первой части теоремы рассмотрим длинные последовательности принятого сообщения M' и соответствующее исходное сообщение M . Количество недостающих данных в принятом сообщении M' составляет (в логарифмической мере) $TH_y(x)$. Поэтому следует передавать каждые T секунд $TH_y(x)$ двоичных цифр. Это может быть сделано с частотой ошибок ϵ в канале с пропускной способностью, равной $H_y(x)$.

Вторая часть теоремы может быть доказана, если заметить, что, во-первых, для каждого дискретных случайных переменных x , y , z

$$H_y(x, z) \geq H_y(x).$$

Левая часть может быть представлена в развернутом виде

$$H_y(z) + H_{yz}(x) \geq H_y(x),$$

$$H_{yz}(x) \geq H_y(x) - H_y(z) \geq H_y(x) - H(z).$$

Если отождествить x с выходным эффектом источника, y — с принимаемым сигналом, а z — с сигналом, посланным по каналу коррекции, то правая часть равна ненадежности за вычетом ско-

рости передачи по каналу коррекции. Если пропускная способность этого канала меньше, чем величина ненадежности, правая часть будет больше нуля и $H_{vz}(x) > 0$. Но это есть неопределенность того, что было передано, когда известен как принятый сигнал, так и сигнал коррекции. Если эта неопределенность больше нуля, частота ошибок не может быть сколь угодно малой.

П р и м е р. Предположим, что в последовательности двоичных чисел ошибки происходят хаотически, вероятность того, что цифра неправильна, равна p , а вероятность того, что она правильна, равна $q=1-p$. Эти ошибки могут быть исправлены, если известно их положение. Таким образом, канал коррекции должен давать только сведения об этих положениях. Это сводится к передаче сообщений источника, который дает двоичные цифры с вероятностью p для 1 (правильно) и q для 0 (неправильно). Необходимая пропускная способность канала коррекции

$$- [p \log p + q \log q]$$

равна ненадежности исходной системы.

Скорость передачи R может быть записана в двух различных формах на основании приведенных выше тождеств. Имеем

$$R = H(x) - H_v(x) = H(y) - H_x(y) = H(x) + H(y) - H(x, y).$$

Первое выражение всегда интерпретируется как количество переданных данных за вычетом неопределенности того, что было передано. Второе выражение измеряет количество принятых данных за вычетом той части, которая обусловлена шумами. Третье выражение есть сумма количества переданных и количества принятых данных за вычетом «энтропии» совместных событий. Таким образом, все три выражения имеют определенное интуитивное значение.

Пропускная способность канала с шумами должна быть максимально возможной скоростью передачи, т. е. скоростью при должном согласовании источника с каналом.

Определим поэтому пропускную способность канала как

$$C = \max [H(x) - H_v(x)],$$

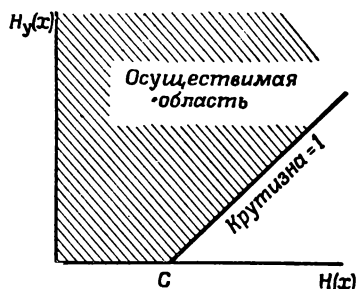
где максимум взят по отношению ко всем возможным источникам сообщений, которые могут быть использованы как входные источники канала. Если канал без шумов, то $H_v(x) = 0$. Тогда это определение эквивалентно тому, которое всегда давалось для канала без шумов, ибо по теореме 8 максимум «энтропии» для канала равен его пропускной способности.

12. ОСНОВНАЯ ТЕОРЕМА ДЛЯ ДИСКРЕТНОГО КАНАЛА ПРИ НАЛИЧИИ ШУМОВ

Может показаться неожиданным, что рассматривается вопрос об определенной пропускной способности канала при наличии шумов, так как в этом случае невозможно передавать сообщения с до-

стоверностью. Однако ясно, что при передаче с избыточными символами вероятность ошибок может быть уменьшена. Например, путем многократного повторения сообщения и путем статистического изучения различных принятых вариантов сообщения вероятность ошибок может быть сделана очень малой. Можно было бы ожидать, что для приближения вероятности ошибок к нулю избыточность кодирования должна неограниченно возрастать, вследствие чего скорость передачи должна приближаться к нулю.

Это ни в коем случае не верно, так как в противном случае не существовало бы вполне определенной пропускной способности канала, а была бы пропускная способность при заданной частоте ошибок или при заданной ненадежности. Тогда пропускная способность уменьшалась бы, по мере того как требования относительно ошибок становились бы более жесткими.



Фиг. 8. Ненадежность, возможная при данной «энтропии» на входе канала.

В действительности пропускная способность имеет вполне определенное значение. При должном кодировании можно передавать по каналу сообщения со скоростью C при сколь угодно малой частоте ошибок или

при сколь угодно малой ненадежности. Это утверждение неверно для скоростей, превышающих C . При попытках передавать со скоростью, превышающей C , скажем $C + R_1$, неизбежно появится ненадежность, равная или большая, чем R_1 .

Это положение иллюстрируется фиг. 8. Скорость создания сообщений в канале отложена по горизонтали, а ненадежность — по вертикали. Любая точка выше жирной линии в заштрихованной области может быть осуществлена, тогда как точки, расположенные ниже жирной линии, не осуществимы. Точки самой линии, вообще говоря, не могут быть получены за исключением обычно двух.

Эти положения являются основным подтверждением правильности предложенного определения C ; они будут сейчас доказаны.

Теорема 11

Пусть дискретный канал обладает пропускной способностью C , а дискретный источник — «энтропией» за 1 сек. H . Если $H < C$, то существует такая система кодирования, что сообщения источника могут быть переданы по каналу с произвольно малой частотой ошибок (или со сколь угодно малой ненадежностью). Если $H > C$, то можно закодировать сообщения источника таким образом, чтобы ненадежность была меньше, чем $H - C + \epsilon$, где ϵ сколь угодно мало.

Не существует способа кодирования, обеспечивающего ненадежность меньшую, чем $H-C$.

Метод доказательства первой части этой теоремы состоит не в указании способа кодирования, обеспечивающего требуемые свойства, а в доказательстве того, что такой код должен существовать в определенной группе кодов. В действительности частота ошибок по этой группе усредняется, и будет показано, что это среднее значение может быть сделано меньше, чем ϵ . Если среднее значение совокупности чисел меньше, чем ϵ , то в ней должно существовать, по крайней мере, одно число, которое меньше ϵ . Это и устанавливает искомый результат.

Пропускная способность канала с шумами была определена как

$$C = \max_x [H(x) - H_y(x)],$$

где x относится ко входу, а y — к выходу. Максимум отыскивается по всем источникам, которые могут быть использованы на входе канала. Если максимум в действительности не достигается ни при каком источнике, то пусть S_0 означает источник, обеспечивающий приблизительно максимальную скорость. Предполагая, что S_0 используется как источник на входе канала, рассмотрим возможные передаваемые и принимаемые последовательности большой длительности T . Можно утверждать следующее:

1. Передаваемые последовательности распадаются на два класса: класс с большой вероятностью, содержащий около $2^{TH(x)}$ членов, и остающиеся последовательности с малой общей вероятностью.

2. Аналогично и принимаемые последовательности распадаются на высоковероятную совокупность приблизительно из $2^{IH(y)}$ членов и маловероятную совокупность из остальных последовательностей.

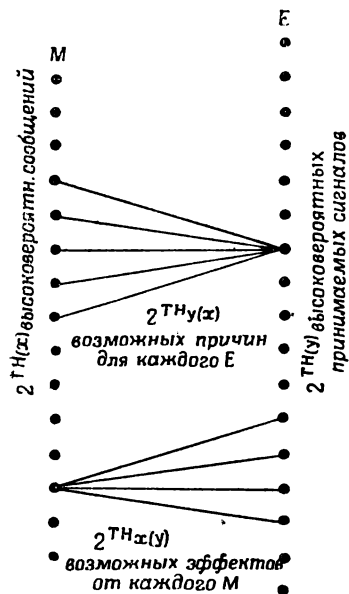
3. Каждый высоковероятный выходной эффект может быть создан примерно $2^{IH_y(x)}$ входными эффектами. Возможность всех остальных случаев имеет малую полную вероятность.

4. Каждый высоковероятный входной эффект может привести примерно к $2^{TH_x(y)}$ выходным эффектам. Все другие результаты имеют малую полную вероятность.

Все ϵ и δ , связанные в этих утверждениях со словами «малый» и «приблизительно», стремятся к нулю, когда T заставляем увеличиваться, а свойства S_0 приближаем к свойствам оптимального источника.

Фиг. 9 иллюстрирует сказанное. На графике входные последовательности представлены точками слева, а выходные — точками справа. Расходящиеся линии («веер») наверху изображают ряд возможных случаев для типичного выходного эффекта. Нижний «веер» представляет возможные случаи для типичного входного эффекта. В обоих примерах отброшены последовательности «малой вероятности».

Предположим теперь, что имеется другой источник, дающий сообщения со скоростью R , причем $R < C$. За период T этот источник будет давать 2^{TR} высоковероятных сообщений. Надо связать их с выбором возможных входных эффектов канала таким образом, чтобы иметь наименьшую частоту ошибок. Будем устанавливать эту связь всеми возможными способами (используя, однако, только высоковероятную группу входных эффектов, определяемых источником S_1) и усредним частоту ошибок для этого широкого класса возможных систем кодирования. Это все равно, что вычислять частоту ошибок для случайной связи сообщений и входных эффектов канала при длительности T .



Фиг. 9. Схематическое представление соотношений между входными и выходными эффектами в канале.

Пусть наблюдается некоторый выходной эффект y_1 . Какова вероятность, что более чем одно сообщение из числа возможных вызовет выходной эффект y_1 ?

Имеются 2^{TR} сообщений, распределенных по случайному закону в $2^{TH(x)}$ точках. Вероятность того, что некоторая данная точка будет сообщением, поэтому равна

$$2^{T[R - H(x)]}.$$

Вероятность того, что ни одна точка «веера» не будет сообщением (кроме действительного исходного сообщения), равна

$$P = (1 - 2^{T[H_y(x) - H(x)]})^{2^{TH_y(x)}}$$

Но $R < [H(x) - H_y(x)]$, так что $R - H(x) = -H_y(x) - \eta$, причем η — положительно. Следовательно:

$$P = (1 - 2^{-TH_y(x) - T\eta})^{2^{TH_y(x)}}$$

при $T \rightarrow \infty$ стремится к

$$1 - 2^{-T\eta}.$$

Отсюда вероятность ошибок стремится к нулю и первая часть теоремы доказана.

Вторую часть теоремы легко доказать, замечая, что можно просто передавать от источника C двоичных единиц в 1 сек., полностью

пренебрегая остатком создаваемых данных. В приемнике пренебрегаемая часть данных создаст ненадежность $H(x) - C$, а передаваемая часть должна лишь добавить ϵ . Этот предел можно также получить многими другими способами, как будет показано при рассмотрении канала с непрерывной передачей.

Последнее утверждение теоремы является прямым следствием определения пропускной способности канала.

Предположим, что можно закодировать сообщения источника, обладающего скоростью $H(x) = C + a$, таким образом, чтобы получить ненадежность $H_y(x) = a - \epsilon$, где ϵ — положительно. Тогда

$$H(x) - H_y(x) = C + \epsilon,$$

где ϵ — положительно. Это противоречит определению C как максимума величины $H(x) - H_y(x)$.

В действительности здесь доказано больше, нежели утверждается в теореме. Если среднее значение множества положительных чисел отличается от нуля меньше, чем на ϵ , то только часть из них, не превышающая $\sqrt{\epsilon}$, может быть больше $\sqrt{\epsilon}$. Так как ϵ сколь угодно мало, можно сказать, что почти все системы сколь угодно близки к идеальной.

13. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Доказательство теоремы 11, не будучи чистым доказательством существования, обладает некоторыми недостатками подобных доказательств. Попытка осуществить хорошее приближение к идеальному кодированию по методу, примененному в доказательстве, вообще говоря, представляется непрактичной. Действительно, за исключением нескольких довольно тривиальных случаев и некоторых предельных положений никаких явных свойств ряда приближений к идеальному методу не найдено. Вероятно, это не случайно, а связано с трудностью задания определенной конструкции, хорошо аппроксимирующей случайную последовательность.

Приближение к идеальному методу обладало бы тем свойством, что оригинал мог бы быть еще восстановлен из сигнала, допустимым образом измененного помехами. Другими словами, предполагается, что это изменение не делает принимаемый сигнал ближе к другим возможным сигналам, чем к оригиналу.

Это достигается ценой введения некоторой избыточности при кодировании. Избыточность должна быть введена соответствующим образом для борьбы против действующих в канале шумов определенной структуры. Всякая избыточность источника будет обычно помогать, если она используется в точке приема. В частности, если источник всегда имеет некоторую избыточность и не принято никаких попыток исключить ее при согласовании с каналом, эта избыточность будет помогать в борьбе с шумами. Например, в телеграфном канале без шумов можно сэкономить около 50% времени путем

правильного кодирования сообщений. Это не делается, и бóльшая часть избыточности английского текста остается в символах канала. В результате при передаче оказываются допустимыми довольно сильные шумы. Значительная часть букв может приниматься неправильно и восстанавливаться на основании контекста.

При отсутствии шумов для приближения к идеальному кодированию требуется, вообще говоря, некоторая временная задержка. Теперь она приобретает новую функцию, позволяя большему «образцу» шумов воздействовать на сигнал, прежде чем какое-либо суждение будет сделано в точке приема относительно исходного сообщения. Увеличение размера «образца» всегда усиливает возможные статистические утверждения.

Содержание теоремы 11 и ее доказательство могут быть сформулированы несколько иным способом, который устанавливает более непосредственную связь со случаем отсутствия шумов. Рассмотрим возможные сигналы длительностью T и предположим, что из них выбрана некоторая подгруппа используемых сигналов. Пусть все сигналы подгруппы употребляются с одинаковой вероятностью; при этом предположим, что приемник устроен так, что, когда принимаются искаженные сигналы, он выбирает в качестве действительного сигнала наиболее вероятный из этой подгруппы. Обозначим через $N(T, q)$ максимальное число сигналов, которые могут быть выбраны для подгруппы таким образом, что вероятность ложной интерпретации меньше или равна q .

Теорема 12

Если C — пропускная способность канала, то при условии, что q не равно 0 или 1

$$\lim_{T \rightarrow \infty} \frac{\log N(T, q)}{T} = C.$$

Другими словами, независимо от требований надежности можно в течение времени T уверенно различить достаточно сообщений, соответствующих примерно CT двоичных единиц, если T достаточно велико. Теорему 12 можно сравнить с определением пропускной способности канала без шумов, данным в главе I.

14. ПРИМЕР ДИСКРЕТНОГО КАНАЛА

Простой пример дискретного канала иллюстрирует фиг. 10. Имеются три возможных символа. Первый символ никогда не подвергается воздействию шумов. Вторым и третьим символами имеют вероятность p пройти неискаженными, и вероятность q превратится в другой символ той же пары. Положим $\alpha = -(p \log p + q \log q)$ и пусть P , Q и Q — вероятности употребления соответственно первого, второго и третьего символов.

Имеем

$$H(x) = -P \lg P - 2Q \lg Q,$$

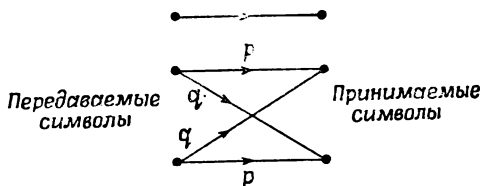
$$H_y(x) = 2Q\alpha.$$

Надо выбрать P и Q таким образом, чтобы $H(x) - H_y(x)$ имело максимум при наложении условия $P + 2Q = 1$. Тогда

$$U = -P \lg P - 2Q \lg Q - 2Q\alpha + \lambda(P + 2Q),$$

$$\frac{\partial U}{\partial P} = -1 - \lg P + \lambda = 0,$$

$$\frac{\partial U}{\partial Q} = -2 - 2 \lg Q - 2\alpha + 2\lambda = 0.$$



Фиг. 10. Пример дискретного канала.

Исключая λ , напишем

$$\lg P = \lg Q + \alpha,$$

$$P = Qe^\alpha = Q\beta,$$

$$P = \frac{\beta}{\beta + 2}; \quad Q = \frac{1}{\beta + 2}$$

Пропускная способность канала равна

$$C = \lg \frac{\beta + 2}{\beta}$$

Заметим, как это подтверждает очевидные значения в случае $\rho = 1$ и $\rho = \frac{1}{2}$. В первом случае $\beta = 1$ и $C = \lg 3$, что правильно, так как канал свободен от шумов и имеет три возможных символа. Если $\rho = \frac{1}{2}$, $\beta = 2$ и $C = \lg 2$. Здесь второй и третий символы не могут быть отличимы друг от друга и действуют совместно как один символ. Первый символ употребляется с вероятностью $P = \frac{1}{2}$, а второй и третий вместе — с вероятностью $\frac{1}{2}$, которая может быть распределена между ними любым способом, причем всегда достигается максимальная пропускная способность.

При промежуточных значениях ρ пропускная способность канала будет заключена между $\lg 2$ и $\lg 3$. Различие между вторым

и третьим символами может быть использовано для передачи некоторого количества сведений, но меньшего, чем в случае канала без шумов. Первый символ употребляется несколько чаще, чем два остальных, так как на него не воздействуют шумы.

15. ПРОПУСКНАЯ СПОСОБНОСТЬ КАНАЛА В НЕКОТОРЫХ СПЕЦИАЛЬНЫХ СЛУЧАЯХ

Если шумы воздействуют на последовательные символы в канале независимо, то для вычисления пропускной способности можно воспользоваться совокупностью вероятностей переходов p_{ij} . Это — вероятность того, что если послан символ i , то будет принят символ j . Максимальная скорость передачи сообщений в канале определяется в этом случае максимумом выражения

$$-\sum_{i,j} P_i p_{ij} \log \sum_i P_i p_{ij} + \sum_{i,j} P_i p_{ij} \log p_{ij},$$

где изменяем P_i при условии $\sum P_i = 1$. По методу Лагранжа это приводит к уравнениям

$$\sum_j p_{sj} \log \frac{p_{sj}}{\sum_i P_i p_{ij}} = \mu, \quad s=1,2,$$

Умножение на P_s и суммирование по s показывают, что $\mu = -C$. Обозначим величины, обратные p_{sj} (если они существуют), через h_{st} , так что $\sum_s h_{st} p_{sj} = b_{it}$.

Тогда

$$\sum_{s,j} h_{st} p_{sj} \log p_{sj} - \log \sum_i P_i p_{it} = -C \sum_s h_{st}.$$

Отсюда

$$\sum_i P_i p_{it} = \exp(C \sum_s h_{st} + \sum_{s,j} h_{st} p_{sj} \log p_{ij})$$

или

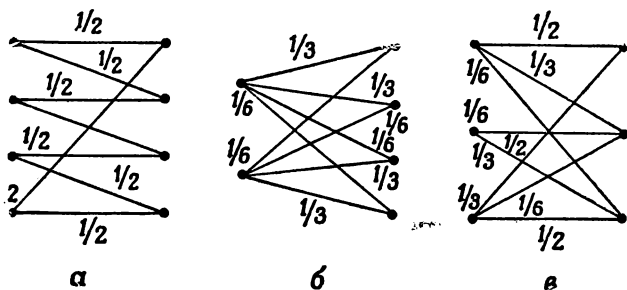
$$P_i = \sum_t h_{it} \exp(C \sum_s h_{st} + \sum_{s,j} h_{st} p_{sj} \log p_{sj})$$

Это — система уравнений для определения оптимальных значений P_i , причем C должно быть определено таким образом, чтобы $\sum P_i = 1$. Когда это сделано, C будет пропускной способностью канала, а P_i — соответствующими вероятностями канальных символов, при которых может быть получена такая пропускная способность.

Если каждый входной символ имеет одинаковую совокупность вероятностей на исходящих от него на диаграмме линиях и то же самое справедливо для каждого выходного символа, пропускная способность может быть легко вычислена.

Примеры показаны на фиг. 11. В таком случае $H_y(x)$ не зависит от распределения вероятностей между входными символами и равняется $-\sum p_i \lg p_i$, где p_i — значения вероятностей переходов от любого входного символа. Пропускная способность канала равна

$$\begin{aligned} \max [H(y) - H_x(y)] &= \\ &= \max H(y) + \sum p_i \log p_i. \end{aligned}$$



Фиг. 11. Примеры дискретных каналов с одинаковыми вероятностями переходов (для каждого входного и выходного эффекта).

Максимум $H(y)$, очевидно, равен $\lg m$, где m — число выходных символов, так как все они могут быть сделаны равновероятными, если сделать равновероятными входные символы. Поэтому пропускная способность канала равна

$$C = \log m + \sum p_i \log p_i.$$

Для случая, приведенного на фиг. 11, а,

$$C = \lg 4 - \log 2 = \log 2.$$

Это значение будет достигнуто при использовании только первого и третьего символов.

Для случая, показанного на фиг. 11, б,

$$\begin{aligned} C &= \log 4 - \frac{2}{3} \log 3 - \frac{1}{3} \log 6 = \\ &= \log 4 - \log 3 - \frac{1}{3} \log 2 = \log \frac{1}{3} 2^{2/3}. \end{aligned}$$

Для случая, приведенного на фиг. 11, в, имеем

$$\begin{aligned} C &= \log 3 - \frac{1}{2} \lg 2 - \frac{1}{3} \log 3 - \frac{1}{6} \log 6 = \\ &= \log \frac{3}{2^{1/2} 3^{1/3} 6^{1/6}}. \end{aligned}$$

Предположим, что символы распадаются на различные группы таким образом, что шумы никогда не превращают символы одной группы в символы другой группы. Пусть пропускная способность для n -й группы равна C_n (в двоичных единицах за 1 сек.), если употребляются только символы этой группы. Тогда легко показать, что для наилучшего использования всей совокупности полная вероятность P_n всех символов в n -й группе должна быть

$$P_n = \frac{2^{C_n}}{\sum 2^{C_n}}.$$

Внутри группы вероятности распределяются как раз так, как если бы эти символы были единственно используемыми. Пропускная способность канала равна

$$C = \lg \sum 2^{C_n}.$$

16. ПРИМЕР ЭФФЕКТИВНОГО КОДИРОВАНИЯ

Следующий пример, хотя и несколько искусственный, является случаем, в котором возможно точное согласование с каналом, подверженным влиянию шумов. Имеются два канальных символа 0 и 1, а шумы воздействуют на них в группах из 7 символов. Группа из 7 символов либо передается без ошибок, либо в ней оказывается ошибочным ровно 1 символ из 7. Эти 8 вероятностей одинаковы. Имеем

$$C = \max [H(y) - H_x(y)] = \frac{1}{7} \left(7 + \frac{8}{8} \lg \frac{1}{8} \right) = \frac{4}{7} \frac{\text{двоичных единиц}}{\text{символ}}$$

Эффективный код, обеспечивающий полную коррекцию ошибок и передачу со скоростью C , представляет собой следующее.

Пусть группа из 7 символов будет X_1, X_2, \dots, X_7 . Из них X_3, X_5, X_6, X_7 — символы сообщения, которые зависят от характера сообщения. Остальные три символа являются избыточными и выбираются следующим образом:

$$\begin{array}{llllll} X_4 & \text{выбирается так, чтобы} & \alpha = X_4 + X_5 + X_6 + X_7 & \text{было четным} & & \\ X_2 & \text{»} & \text{»} & \beta = X_2 + X_3 + X_6 + X_7 & \text{»} & \text{»} \\ X_1 & \text{»} & \text{»} & \gamma = X_1 + X_3 + X_5 + X_7 & \text{»} & \text{»} \end{array}$$

Когда группа из 7 символов принята, вычисляются α, β, γ , и если они окажутся четными, то означают нуль, а если нечетными, то означают единицу. Двоичные цифры α, β, γ дадут тогда индексы тех X_i , которые являются ошибочными (если получится 0, то это означает отсутствие ошибок).

Г л а в а III

НЕПРЕРЫВНЫЕ СООБЩЕНИЯ

Рассмотрим теперь случай, когда сигналы или сообщения (или те и другие) являются непрерывными переменными в противоположность исследованным ранее дискретным системам. При этом значительная часть результатов может быть получена предельным переходом от дискретного случая путем деления всего континуума сообщений или сигналов на большое, но конечное число малых областей и вычисления различных параметров, введенных на дискретной основе. По мере уменьшения размеров областей эти параметры в общем сходятся в пределе к соответствующим значениям для непрерывного случая.

В непрерывном случае не будем стремиться к наибольшей общности или к полной математической строгости, так как это связано с широким применением абстрактной теории размерностей. Предварительное изучение, однако, показывает, что теория может быть сформулирована совершенно аксиоматическим и строгим образом, включая как непрерывный и дискретный случаи, так и многие другие. Некоторые вольности, допущенные в настоящем анализе при предельных переходах, во всех случаях, представляющих практический интерес, могут быть оправданы.

17. МНОЖЕСТВА И АНСАМБЛИ ФУНКЦИЙ

В непрерывном случае встречаемся со множествами функций и с ансамблями функций. Множество функций, как указывает само название, есть просто некоторый класс или набор функций обычно одной переменной — времени. Оно может быть определено либо путем явного представления различных функций во множестве, либо неявно, путем указания тех свойств, которыми обладают функции множества, а другие функции нет. Приводим некоторые примеры:

1. Множество функций

$$f_{\theta}(t) = \sin(t + \theta).$$

Каждое частное значение θ определяет частную функцию множества.

2. Множество всех функций времени, не содержащих частот выше W гц.

3. Множество всех функций, ограниченных по полосе частотой W и по амплитуде — амплитудой A .

4. Множество всех английских речевых сигналов, рассматриваемых как функции времени.

Ансамбль функций есть множество функций вместе с вероятностной мерой, посредством которой можно определить вероятность того, что функция множества обладает некоторыми определенными свойствами ¹⁾. Например, вместе со множеством

$$f_{\theta}(t) = \sin(t + \theta)$$

можно задать распределение вероятностей для θ , т. е. $P(\theta)$. В таком случае множество становится ансамблем.

Приведем некоторые другие примеры ансамблей функций:

1. Конечное множество функций $f_k(t)$, где $k=1, 2, \dots, n$, вместе с вероятностью того, что f_k есть p_k .

2. Множество функций с конечным числом измерений

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

вместе с распределением вероятностей для параметров α_i

$$p(\alpha_1, \alpha_2, \dots, \alpha_n).$$

Например, можно рассмотреть ансамбль, определяемый в виде

$$f(a_1, \dots, a_n, \theta_1, \dots, \theta_n; t) = \sum_{n=1}^n a_n \sin n(\omega t + \theta_n),$$

где амплитуды a_i распределены нормально и независимо, а фазы θ_i распределены равномерно и независимо в интервале $(0 \div 2\pi)$.

3. Ансамбль

$$f(a_i, t) = \sum_{n=-\infty}^{+\infty} a_n \frac{\sin \pi(2Wt - n)}{\pi(2Wt - n)},$$

где a_i распределены по нормальному закону, независимы и все имеют одно и то же стандартное отклонение \sqrt{N} . Это выражение представляет «белые» шумы, полоса частот которых ограничена участком от 0 до W гц, а средняя мощность равна N^2 .

¹⁾ По математической терминологии функции принадлежат к измеримому пространству, полная мера которого есть единица.

²⁾ Это представление может быть использовано как определение «белых» шумов с ограниченной полосой частот. Оно имеет некоторые преимущества, связанные с меньшим числом предельных переходов, нежели определения, применявшиеся в прошлом. Термин «белые шумы», прочно укоренившийся в литературе, представляется несколько неудачным. В оптике под белым светом понимается излучение, имеющее либо сплошной спектр (в противоположность линейчатому), либо спектр, равномерный по отношению к длине волны (а это не то же самое, что спектр, равномерный по отношению к частоте).

4. Пусть на оси t распределены точки по закону Пуассона. В каждой избранной точке помещается функция $f(t)$ и различные функции складываются, давая ансамбль

$$\sum_{k=-\infty}^{\infty} f(t + t_k),$$

где t_k — точки, подчиняющиеся распределению Пуассона. Этот ансамбль может рассматриваться как разновидность импульсных или дробовых шумов, когда все импульсы одинаковы.

5. Система английских речевых функций с вероятностной мерой, определяемой частотой повторения при обычном использовании.

Ансамбль функций $f_a(t)$ называется *стационарным*, если при сдвиге всех функций во времени на некоторую фиксированную величину получается тот же ансамбль.

Например, ансамбль

$$f_{\theta}(t) = \sin(t + \theta)$$

является стационарным, если θ равномерно распределены в интервале $0 \div 2\pi$. Если сдвинуть каждую функцию на t_1 , то получается

$$f_{\theta}(t + t_1) = \sin(t + t_1 + \theta) = \sin(t + \varphi),$$

где φ распределены равномерно в интервале $0 \div 2\pi$. Каждая функция изменилась, но ансамбль в целом при этом смещении остался неизменным. В приведенных выше других примерах ансамбли также все стационарны.

Ансамбль называется *эргодическим*, если он является стационарным и если во множестве функций не существует подмножества функций с вероятностью, отличной от 0 и 1, которое было бы стационарным. Ансамбль

$$\sin(t + \theta)$$

является эргодическим. Никакое подмножество этих функций с вероятностью, отличной от 0 и 1, не может быть превращено в самое себя при всех временных смещениях.

Вместе с тем ансамбль

$$a \sin(t + \theta),$$

где a распределены по нормальному закону, а θ — равномерно, является стационарным, но не эргодическим. Подмножество этих функций для a , заключенных между 0 и 1, например, является стационарным и имеет вероятность, не равную 0 или 1.

Из приведенных выше примеров ансамблей 3-й и 4-й являются эргодическими, а 5-й, возможно, также может рассматриваться как эргодический. Если ансамбль эргодический, то, грубо говоря, каждая функция множества является типичной для ансамбля.

Более точно, в случае эргодического ансамбля среднее любой статистической величины, взятое по ансамблю, равно (с вероятностью 1) среднему, взятому по всем временным смещениям некоторой частной функции множества¹⁾. Грубо говоря, можно ожидать, что по мере течения времени: каждая функция испытает с надлежащей частотой все изменения, претерпеваемые любой из функций множества.

Выполняя различные операции с числами или функциями, получаем новые числа или функции. Точно так же можно совершать операции над ансамблями для получения новых ансамблей. Допустим, например, что имеется ансамбль функций $f_a(t)$ и оператор T , который даст для каждой функции $f_a(t)$ результат

$$g_a(t) = T f_a(t).$$

Для множества $g_a(t)$ мера вероятностей определяется мерой для множества $f_a(t)$. Вероятность некоторого подмножества функций $g_a(t)$ равна вероятности подмножества функций $f_a(t)$, которые создают члены данного подмножества функций g в результате операции T . Физически это соответствует прохождению ансамбля через некоторое устройство, например фильтр, выпрямитель или модулятор. Функции на выходе устройства образуют ансамбль $g_a(t)$.

Устройство или оператор будет называться инвариантным, если сдвиг входной функции приводит просто к сдвигу выходной функции, т. е. если

$$g_a(t) = T f_a(t)$$

означает, что

$$g_a(t+t_1) = T f_a(t+t_1)$$

для всех $f_a(t)$ и всех t_1 . Легко показать (см. Приложение 5), что если T — инвариантный оператор, а входной ансамбль стационарный, то выходной ансамбль также стационарный. Подобным же образом, если входной ансамбль эргодический, то выходной ансамбль будет также эргодическим.

Фильтр или выпрямитель являются инвариантными устройствами при всех временных смещениях. Операция модуляции не является инвариантной, так как фаза несущей создает определенную временную структуру. Однако модуляция инвариантна при всех смещениях, кратных периоду несущей.

Существует тесная связь между инвариантностью физических устройств при временных смещениях и теорией Фурье. Если устройство линейно и инвариантно, то анализ методами Фурье является удобным математическим аппаратом для решения задачи.

¹⁾ Это и есть знаменитая эргодическая теорема Биркгофа—Хинчина или, вернее, один из вариантов этой теоремы. См., например, Э. Хопф, «Эргодическая теория», *Успехи математических наук*, IV, вып. 1, 1949. (Прим. ред.)

Ансамбль функций представляет собой подходящее математическое представление для сообщений, создаваемых непрерывным источником (например, речь), для сигналов от передатчика и для мешающих шумов. Теория связи имеет дело не с операциями над частными функциями, а с операциями над ансамблями функций. Система связи конструируется не для определенной речевой функции и тем более не для синусоидальной функции, а для ансамбля речевых функций.

18. АНСАМБЛИ ФУНКЦИЙ С ОГРАНИЧЕННОЙ ПОЛОСОЙ ЧАСТОТ

Если функция времени $f(t)$ ограничена по полосе частот участком от 0 до W гц, то она полностью определяется заданием ее ординат в ряде дискретных точек, отстоящих друг от друга на $\frac{1}{2W}$ сек.

Теорема 13¹⁾

Пусть $f(t)$ не содержит частот, превышающих W гц.

Тогда

$$f(t) = \sum_{-\infty}^{\infty} X_n \frac{\sin \pi(2Wt - n)}{\pi(2Wt - n)},$$

где

$$X_n = f\left(\frac{n}{2W}\right)$$

В этом разложении $f(t)$ представлена как сумма ортогональных функций. Коэффициенты X_n при различных членах могут рассматриваться как координаты в «функциональном пространстве» с бесконечным числом измерений. В этом пространстве каждой функции соответствует одна точка и каждой точке — одна функция.

Функция может рассматриваться как ограниченная временным интервалом T , если все ординаты X_n вне этого временного интервала будут равны нулю. В этом случае только $2TW$ координат отличаются от нуля. Таким образом, функции, ограниченные полосой частот W и длительностью T , соответствуют точкам в пространстве $2TW$ измерений.

Подмножество функций с полосой частот W и длительностью T соответствует области в этом пространстве. Например, функции, полная энергия которых меньше или равна E , соответствуют точкам сферы из $2TW$ измерений с радиусом $r = \sqrt{2WE}$.

Ансамбль функций с ограниченной полосой частот и ограниченной длительностью будет представляться распределением

¹⁾ Эта теорема была установлена В. А. Котельниковым в 1933 г. (Прим. ред.)

вероятностей $p(x_1, \dots, x_n)$ в соответствующем n -мерном пространстве. Если ансамбль не ограничен по времени, то можно считать, что $2TW$ координат в данном интервале T представляют часть функции в интервале T , а распределение вероятностей $p(x_1, \dots, x_n)$ дает статистическую структуру ансамбля для интервалов такой длительности.

19. «ЭНТРОПИЯ» НЕПРЕРЫВНОГО РАСПРЕДЕЛЕНИЯ

«Энтропия» дискретной группы вероятностей p_1, \dots, p_n определялась как

$$H = - \sum p_i \log p_i.$$

Аналогичным образом определим «энтропию» непрерывного распределения с функцией плотности распределения $p(x)$ как

$$H = - \int_{-\infty}^{\infty} p(x) \log p(x) dx.$$

В случае n -мерного распределения $p(x_1, \dots, x_n)$ имеем

$$H = - \int \dots \int p(x_1, \dots, x_n) \log p(x_1, \dots, x_n) dx_1 \dots dx_n.$$

Если имеются два аргумента x и y (которые сами могут быть многомерными), то «энтропия» совместного события и условная «энтропия» даются уравнениями

$$H(x, y) = - \iint p(x, y) \log p(x, y) dx dy,$$

$$H_x(y) = - \iint p(x, y) \log \frac{p(x, y)}{p(x)} dx dy,$$

$$H_y(x) = - \iint p(x, y) \log \frac{p(x, y)}{p(y)} dx dy,$$

где

$$p(x) = \int p(x, y) dy,$$

$$p(y) = \int p(x, y) dx.$$

«Энтропия» непрерывного распределения имеет многие свойства дискретного случая. В частности:

1. Если x ограничено некоторым объемом v в своем пространстве, то $H(x)$ максимальна и равна $\log v$, когда $p(x)$ постоянно и равно $\frac{1}{v}$ в этом объеме.

2. При любых двух переменных x, y имеем

$$H(x, y) \leq H(x) + H(y),$$

причем знак равенства будет тогда (и только тогда), когда x и y независимы, т. е. $p(x, y) = p(x)p(y)$ (за исключением, возможно, ряда точек с нулевой вероятностью).

3. Рассмотрим обобщенную операцию усреднения следующего типа:

$$p'(y) = \int a(x, y)p(x) dx,$$

где

$$\int a(x, y) dx = \int a(x, y) dy = 1, \quad a(x, y) \geq 0.$$

Тогда «энтропия» усредненного распределения $p'(y)$ равна или больше «энтропии» первоначального распределения $p(x)$.

4. Имеем

$$H(x, y) = H(x) + H_x(y) = H(y) + H_y(x)$$

и

$$H_x(y) \leq H(y).$$

5. Пусть $p(x)$ будет одномерное распределение. Распределение $p(x)$, обеспечивающее максимальную «энтропию», при условии, что стандартное отклонение x равно σ , есть нормальное распределение. Чтобы это показать, необходимо разыскать максимум

$$H(x) = - \int p(x) \log p(x) dx,$$

накладывая в качестве ограничений соотношение

$$\sigma^2 = \int p(x)x^2 dx \text{ и } 1 = \int p(x) dx.$$

Согласно вариационному исчислению это приводит к необходимости нахождения максимума для

$$\int [-p(x) \log p(x) + \lambda p(x)x^2 + \mu p(x)] dx.$$

Условием этого является

$$-1 - \log p(x) + \lambda x^2 + \mu = 0,$$

и, следовательно, подбирая постоянные для удовлетворения упомянутых ограничений, найдем

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$$

Подобным же образом обстоит дело и в случае n измерений, когда моменты второго порядка распределения $p(x_1, \dots, x_n)$:

$$A_{ij} = \int \dots \int x_i x_j p(x_1, \dots, x_n) dx_1 \dots dx_n.$$

При помощи подобных же вычислений найдем, что максимум «энтропии» имеет место, когда $p(x_1, \dots, x_n)$ является n -мерным нормальным распределением с моментами второго порядка A_{ij} .

6. «Энтропия» одномерного нормального распределения, стандартное отклонение которого есть σ , равна

$$H(x) = \log \sqrt{2\pi e} \sigma.$$

Ход вычислений таков:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \quad -\log p(x) = \log \sqrt{2\pi}\sigma + \frac{x^2}{2\sigma^2},$$

$$\begin{aligned} H(x) &= -\int p(x) \log p(x) dx = \int p(x) \log \sqrt{2\pi}\sigma dx + \int p(x) \frac{x^2}{2\sigma^2} dx = \\ &= \log \sqrt{2\pi}\sigma + \frac{\sigma^2}{2\sigma^2} = \log \sqrt{2\pi}\sigma + \log \sqrt{e} = \log \sqrt{2\pi e} \sigma. \end{aligned}$$

Подобным же образом n -мерное нормальное распределение с квадратичной формой a_{ij} дается уравнением

$$p(x_1, \dots, x_n) = \frac{|a_{ij}|^{1/2}}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2} \sum a_{ij} X_i X_j\right),$$

а «энтропия» может быть вычислена как

$$H = \log (2\pi e)^{n/2} |a_{ij}|^{-1/2},$$

где $|a_{ij}|$ есть определитель, элементы которого a_{ij} .

7. Если x ограничено половиной оси [$p(x)=0$ при $x \leq 0$], а первый момент x равен

$$a = \int_0^{\infty} p(x) x dx,$$

то максимум «энтропии» имеет место, когда

$$p(x) = \frac{1}{a} e^{-\frac{x}{a}}$$

и равен он

$$\log ea.$$

8. Между «энтропиями» для непрерывного и дискретного случаев имеется одно существенное различие. В дискретном случае «энтропия» измеряет абсолютным образом беспорядочный характер случайной переменной. В непрерывном случае измерение является относительным к координатной системе. Если изменить координаты, то «энтропия» в общем случае также изменится. Действительно, при переходе к координатам y_1, \dots, y_n новое значение «энтропии» будет

$$H(y) = \int \dots \int p(x_1, \dots, x_n) J\left(\frac{x}{y}\right) \log p(x_1, \dots, x_n) J\left(\frac{x}{y}\right) dy_1 \dots dy_n,$$

где $J\left(\frac{x}{y}\right)$ есть якобиан преобразования координат. Разлагая логарифм и меняя переменные на x_1, \dots, x_n , получим

$$H(y) = H(x) - \int \dots \int p(x_1, \dots, x_n) \log J\left(\frac{x}{y}\right) dx_1 \dots dx_n.$$

Таким образом, новое значение «энтропии» равно старому за вычетом ожидаемого логарифма якобиана. В непрерывном случае «энтропия» может рассматриваться как мера случайности *относительно принятого стандарта*, а именно выбранной координатной системы, в которой каждому малому элементу объема $dx_1 \dots dx_n$ придан равный вес. При изменении координатной системы «энтропия» в новой системе является мерой случайности, когда равным элементам объема $dy_1 \dots dy_n$ в новой системе придан одинаковый вес.

Несмотря на зависимость от координатной системы, понятие «энтропии» является столь же важным в непрерывном случае, как и в дискретном. Это объясняется тем, что скорость создания сообщений и пропускная способность канала определяются *разностью* двух «энтропий», а эта разность *не зависит* от координатной системы, так как каждый из двух членов изменяется одинаково.

«Энтропия» непрерывного распределения может быть отрицательной. Шкала измерений устанавливает произвольный нуль, соответствующий равномерному распределению по единичному объему. Распределение, более сосредоточенное чем это, будет иметь меньшую «энтропию», и следовательно, она отрицательна. Однако скорость создания сообщений и пропускная способность канала всегда будут не отрицательны.

9. Частным случаем изменения координат является линейное преобразование

$$y_j = \sum_i a_{ij} x_i.$$

При этом якобиан есть просто определитель $|a_{ij}|^{-1}$ и

$$H(y) = H(x) + \log |a_{ij}|.$$

В случае вращения координатной системы (или любого другого преобразования, сохраняющего измерение) $J=1$ и $H(y)=H(x)$.

20. «ЭНТРОПИЯ» АНСАМБЛЯ ФУНКЦИЙ

Рассмотрим эргодический ансамбль функций с ограниченной полосой частот шириной W гц. Пусть

$$p(x_1, \dots, x_n)$$

будет функция плотности распределения для амплитуд x_1, \dots, x_n в n последовательных точках. Определим «энтропию» ансамбля функций на степень свободы как

$$H' = - \lim_{n \rightarrow \infty} \frac{1}{n} \int \int p(x_1, \dots, x_n) \log p(x_1, \dots, x_n) dx_1 \dots dx_n.$$

Можно также определить «энтропию» H за 1 сек. путем деления не на n , а на время T в секундах для n значений. Так как $n=2TW$, то $H = 2WH'$.

Для «белых» термических шумов распределение p является нормальным и поэтому

$$H' = \log \sqrt{2\pi eN},$$

$$H = W \log 2\pi eN.$$

При данной средней мощности N «белые» шумы имеют максимально возможную «энтропию». Это следует из отмеченных выше максимальных свойств нормального распределения.

«Энтропия» непрерывного стохастического процесса имеет много свойств, аналогичных свойствам энтропии дискретных процессов. В дискретном случае «энтропия» была связана с логарифмом вероятности длительных последовательностей и с числом сравнительно вероятных последовательностей большой длительности. В непрерывном случае «энтропия» подобным же образом связана с логарифмом плотности вероятностей для длинной серии «образцов» сигнала и с объемом сравнительно высокой вероятности в функциональном пространстве.

Более точно, если положить, что $p(x_1, \dots, x_n)$ является непрерывным при всех x_i и для всех n , то для сравнительно больших n

$$\left| \frac{\log p}{n} - H' \right| < \epsilon$$

при любом выборе (x_1, \dots, x_n) за исключением группы, полная вероятность которой меньше δ , причем δ и ϵ произвольно малы. Это следует из эргодических свойств при разделении пространства на большое число малых ячеек.

Связь H с объемом может быть установлена следующим образом. При тех же самых предположениях рассмотрим n -мерное пространство, соответствующее $p(x_1, \dots, x_n)$. Пусть $V_n(q)$ будет наименьший объем в этом пространстве, который заключает в себе полную вероятность q . Тогда

$$\lim_{n \rightarrow \infty} \frac{\log V_n(q)}{n} = H',$$

если только q не равно 0 или 1.

Эти результаты показывают, что при больших n существует совершенно четко определенный объем высокой вероятности (по крайней мере, в логарифмическом смысле) и что внутри этого объема плотность вероятностей сравнительно равномерна (опять-таки в логарифмическом смысле).

В случае «белых» шумов плотность вероятностей равна

$$p(x_1, \dots, x_n) = \frac{1}{(2\pi N)^{n/2}} \exp\left(-\frac{1}{2N} \sum x_i^2\right).$$

Так как она зависит только от $\sum x_i^2$, то поверхности равной плотности вероятностей представляют собой сферы, а все распределение обладает сферической симметрией. Область высокой вероятности есть сфера радиуса \sqrt{nN} . При $n \rightarrow \infty$ вероятность нахождения вне сферы радиуса $\sqrt{n(N+\epsilon)}$ стремится к нулю, как бы ни было мало ϵ , а $\frac{1}{n}$ логарифма объема сферы стремится к $\log \sqrt{2\pi e N}$.

В непрерывном случае удобно пользоваться не «энтропией» ансамбля H , а производной величиной, которую будем называть *энтропийной мощностью*. Она определяется как мощность «белых» шумов, ограниченных такой же полосой частот, что и рассматриваемый ансамбль, и имеющих такую же «энтропию». Другими словами, если H' есть «энтропия» ансамбля, то его энтропийная мощность равна

$$N_1 = \frac{1}{2\pi e} e^{2H'}.$$

В геометрической трактовке это означает измерение объема высокой вероятности квадратом радиуса сферы, имеющей такой же объем. Так как «белые» шумы имеют максимальную «энтропию» для данной мощности, то энтропийная мощность любых шумов меньше или равна их действительной мощности.

21. ПОТЕРЯ «ЭНТРОПИИ» В ЛИНЕЙНЫХ ФИЛЬТРАХ

Т е о р е м а 14

Если ансамбль функций, имеющих «энтропию» на степень свободы H_1 в полосе частот W , пропускается через фильтр с характеристикой $Y(f)$, то ансамбль на выходе имеет «энтропию»

$$H_2 = H_1 + \frac{1}{W} \int \log |Y(f)|^2 df.$$

Действие фильтра представляет собой линейное преобразование координат. Если рассматривать частотные составляющие как первоначальные координаты системы, то новые частотные состав-

ляющие будут представлять собой старые, умноженные на некоторые коэффициенты. Таким образом, матрица преобразования координат является относительно этих координат диагональной. Якобиан преобразования равен (для n синусоидальных и n косинусоидальных составляющих)

$$J = \prod_{i=1}^n |Y(f_i)|^2 = \exp \sum \log |Y(f_i)|^2,$$

где f_i расположены на равных расстояниях в полосе W . В пределе это превращается в

$$\exp \frac{1}{W} \int_W \log |Y(f)|^2 df.$$

Так как якобиан J постоянен, то его среднее значение равно такой же величине. Применяя теорему об изменении «энтропии» с изменением координат, получаем сформулированный выше результат. Его можно также выразить через энтропийную мощность.

Поэтому если энтропийная мощность первого ансамбля функций есть N_1 , то энтропийная мощность второго ансамбля равна

$$N_1 \exp \frac{1}{W} \int_W \log |Y(f)|^2 df.$$

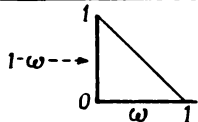
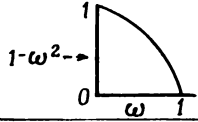
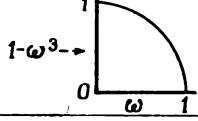
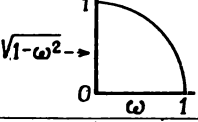
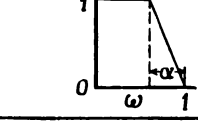
Конечная энтропийная мощность равна начальной, умноженной на геометрическое среднее коэффициента передачи фильтра. Если усиление измеряется в децибелах, то выходная энтропийная мощность увеличится на арифметическое среднее коэффициента передачи в децибелах внутри полосы W .

В табл. 1 для ряда идеализированных характеристик передачи фильтра вычислена потеря энтропийной мощности, выраженная в децибелах. Приведены тоже импульсные характеристики этих фильтров для $W=2\pi$, причем предполагается, что фазовый угол равен нулю.

Потеря «энтропии» во многих других случаях может быть найдена при помощи этих результатов. Например, коэффициент энтропийной мощности $\frac{1}{e^2}$, полученный для первого случая, относится также к любой характеристике передачи, получаемой из $1-\omega$ путем преобразования оси ω с сохранением измерения. В частности, линейно возрастающая характеристика $G(\omega)=\omega$ или пилообразная характеристика между 0 и 1 имеет такую же потерю «энтропии». Обратная характеристика имеет обратный коэффициент потерь,

поэтому в случае характеристики $\frac{1}{\omega}$ коэффициент равен e^2 . Возведение усиления в какую-либо степень приводит к возведению коэффициента потерь в ту же степень.

Таблица 1

Характеристика фильтра	Коэффициент энтропной мощности	Усиление энтропной мощности (дб)	Импульсная реакция
	$\frac{1}{e^2}$	-8,68	$\frac{\sin^2 \pi t}{(\pi t)^2}$
	$\left(\frac{2}{e}\right)^4$	-5,32	$2\left(\frac{\sin t}{t^3} - \frac{\cos t}{t^2}\right)$
	0,384	-4,15	$6\left(\frac{\cos t - 1}{t^4} - \frac{\cos t}{2t^2} + \frac{\sin t}{t^3}\right)$
	$\left(\frac{2}{e}\right)^2$	-2,66	$\frac{\pi}{2} \frac{J_1(t)}{t}$
	$\frac{1}{e^{2\alpha}}$	-8,68 alpha	$\frac{1}{\alpha t^2} [\cos(1-\alpha)t - \cos t]$

22. «ЭНТРОПИЯ» СУММЫ ДВУХ АНСАМБЛЕЙ ФУНКЦИЙ

Если имеются два ансамбля функций $f_a(t)$ и $g_b(t)$, то можно создать новый ансамбль путем «сложения». Допустим, что первый имеет функцию плотности вероятностей $p(x_1, \dots, x_n)$, а второй — $q(x_1, \dots, x_n)$. Тогда функция плотности для суммы дается выражением

$$r(x_1, \dots, x_n) = \int \dots \int p(y_1, \dots, y_n) q(x_1 - y_1, \dots, x_n - y_n) dy_1 \dots dy_n.$$

Физически это соответствует сложению шумов или сигналов, представляемых первоначальными ансамблями функций.

Следующий результат доказывается в Приложении 6.

Теорема 15

Пусть средняя мощность двух ансамблей функций будет N_1 и N_2 , а их энтропийные мощности — \bar{N}_1 и \bar{N}_2 . Тогда энтропийная мощность суммы \bar{N}_3 ограничена пределами

$$\bar{N}_1 + \bar{N}_2 \leq \bar{N}_3 \leq N_1 + N_2.$$

«Белые» шумы с нормальным распределением имеют свойство поглощать всякие другие шумы или ансамбли сигналов, которые могут быть сложены с ними. При этом результирующая энтропийная мощность приближенно равна сумме мощности «белых» шумов и мощности сигнала (измеренной от среднего значения сигнала, которое обычно равно нулю), если только мощность сигнала мала (в определенном смысле) по сравнению с шумами.

Рассмотрим функциональное пространство n измерений, связанное с этими ансамблями функций. «Белые» шумы соответствуют сферическому нормальному распределению в этом пространстве. Ансамбль сигналов соответствует другому распределению, не обязательно нормальному или сферическому.

Пусть моменты второго порядка этого распределения относительно его центра тяжести будут a_{ij} . Другими словами, если $p(x_1, \dots, x_n)$ есть функция плотности распределения, то

$$a_{ij} = \int \dots \int p(x_i - \alpha_i)(x_j - \alpha_j) dx_1 \dots dx_n,$$

где α_i координаты центра тяжести, а a_{ij} — определенно положительная квадратичная форма. Повернув координатную систему, можно выравнять ее с главными направлениями этой формы. Тогда a_{ij} приводится к диагональной форме b_{ii} . Потребуем, чтобы каждая форма b_{ii} была мала сравнительно с N — квадратом радиуса сферического распределения.

В этом случае шумы и сигнал создают нормальное распределение, соответствующая квадратичная форма которого есть

$$N + b_{ii}.$$

Энтропийная мощность этого распределения равна

$$[\prod (N + b_{ii})]^{1/n}$$

или приближенно

$$(N^n + \sum b_{ii} N^{n-1})^{1/n} = N + \frac{1}{n} \sum b_{ii}.$$

Последний член есть мощность сигнала, первый — мощность шумов.

Глава IV

КАНАЛ С НЕПРЕРЫВНОЙ ПЕРЕДАЧЕЙ

23. ПРОПУСКНАЯ СПОСОБНОСТЬ КАНАЛА С НЕПРЕРЫВНОЙ ПЕРЕДАЧЕЙ

В канале с непрерывной передачей входные или передаваемые сигналы являются непрерывными функциями времени $f(t)$, принадлежащими к некоторому множеству, а выходные или принимаемые сигналы будут их искаженными вариантами. Рассмотрим только такой случай, когда как передаваемые, так и принимаемые сигналы ограничены некоторой полосой частот W . Тогда в интервале T они могут быть заданы $2TW$ числами, а их статистическая структура описана функциями распределения с конечным числом измерений. Таким образом, статистические свойства передаваемого сигнала будут определяться функцией

$$P(x_1, \dots, x_n) = P(x),$$

а статистические свойства шумов — распределением условных вероятностей

$$P_{x_1, \dots, x_n}(y_1, \dots, y_n) = P_x(y).$$

Скорость передачи сообщений в таком канале определяется аналогично дискретному каналу, а именно:

$$R = H(x)_i - H_y(x),$$

где $H(x)$ есть «энтропия» на входе, а $H_y(x)$ — ненадежность.

Пропускная способность канала C определяется как максимум R при изменении входного ансамбля по всем возможным ансамблям. Это означает, что в случае приближения с конечным числом измерений следует изменять $P(x) = P(x_1, \dots, x_n)$ и разыскивать максимум

$$-\int P(x) \log P(x) dx + \iint P(x, y) \log \frac{P(x, y)}{P(y)} dx dy.$$

Это выражение может быть написано в виде

$$\iint P(x, y) \log \frac{P(x, y)}{P(x) P(y)} dx dy,$$

пользуясь тем обстоятельством, что

$$\iint P(x, y) \log P(x) dx dy = \int P(x) \log P(x) dx.$$

Таким образом, пропускная способность канала выражается как

$$C = \lim_{T \rightarrow \infty} \max_{P(x)} \frac{1}{T} \iint P(x, y) \log \frac{P(x, y)}{P(x)P(y)} dx dy.$$

Отсюда ясно, что R и C не зависят от координатной системы, так как числитель и знаменатель в $\log \frac{P(x, y)}{P(x)P(y)}$ при преобразовании x и y любым однозначным образом будут умножаться на один и тот же коэффициент.

Интегральное выражение для C является более общим, чем выражение $H(x) - H_y(x)$. Надлежащим образом интерпретированное, оно всегда существует, тогда как в некоторых случаях $H(x) - H_y(x)$ может оказаться неопределенной формой вида $\infty - \infty$. Это происходит, например, если в случае n -мерной аппроксимации x ограничивается поверхностью меньшего числа измерений, нежели n .

Если используемое при вычислении $H(x)$ и $H_y(x)$ основание логарифмов равно двум, то, как и в дискретном случае, C есть максимальное число двоичных единиц, которое может быть передано за 1 сек. по каналу со сколь угодно малой ненадежностью. Это можно понять физически, разделив пространство сигналов на большое число маяжк ячеек. Ячейки делаются настолько малыми, чтобы плотность вероятностей того, что сигнал x в результате действия шумов перейдет в точку y , т. е. $P_x(y)$, была достаточно постоянной по всей ячейке. Если ячейки рассматриваются как отдельные точки, то положение будет точно такое же, что и в дискретном канале, и использованные там доказательства будут применимы и здесь.

Физически ясно, что подразделение объема на отдельные точки в любых практических случаях не может существенно сказаться на конечном результате, если только ячейки достаточно малы. Поэтому пропускная способность будет пределом пропускных способностей для дискретных подразделений, а это и есть пропускная способность канала с непрерывной передачей, как она определена выше.

Математически можно прежде всего показать, что если u есть сообщение, x — сигнал, y — принимаемый сигнал (измененный шумами), а v — восстановленное из сигнала сообщение, то

$$H(x) - H_y(x) \geq H(u) - H_v(u)$$

независимо от того, какие операции производились над u , чтобы получить x , или над y для получения v .

Таким образом, независимо от того, как кодируют двоичные знаки для создания сигнала или как декодируют принимаемый сигнал для восстановления сообщения, скорость дискретной передачи двоичных знаков не превышает определенную выше пропускную способность канала. С другой стороны, при весьма общих условиях

можно найти систему кодирования, обеспечивающую передачу двоичных знаков со скоростью C при сколь угодно малой ненадежности или частоте ошибок. Это справедливо, например, когда в случае аппроксимации сигнальных функций пространством конечного числа измерений, $P(x, y)$ непрерывно как по x , так и по y , за исключением группы точек, где вероятность равна нулю.

Важный частный случай имеет место, когда шумы складываются с сигналом, являясь независимыми от него (в вероятностном смысле). Тогда $P_x(y)$ есть функция только разности (векторной) $n = y - x$

$$P_x(y) = Q(y - x)$$

и шумам можно приписать определенную «энтропию» (независимо от статистических свойств сигнала), а именно «энтропию» распределения $Q(n)$. Эта «энтропия» будет обозначаться $H(n)$.

Теорема 16

Если сигнал и шумы независимы, а принимаемый сигнал является суммой передаваемого сигнала и шумов, то скорость передачи равна

$$R = H(y) - H(n),$$

т. е. «энтропии» принимаемого сигнала за вычетом «энтропии» шумов. Пропускная способность канала равна

$$C = \max_{P(x)} H(y) - H(n).$$

В силу того, что $y = x + n$, имеем

$$H(x, y) = H(x, n).$$

Разлагая левую часть и пользуясь независимостью x и n , найдем

$$H(y) + H_y(x) = H(x) + H(n).$$

Отсюда

$$R = H(x) - H_y(x) = H(y) - H(n).$$

Так как $H(n)$ не зависит от $P(x)$, то для максимума R необходимо, чтобы имела максимум $H(y)$ — «энтропия» принимаемого сигнала. Если на ансамбль передаваемых сигналов накладываются некоторые ограничения, то «энтропия» принимаемых сигналов должна быть максимальной при этих ограничениях.

24. ПРОПУСКНАЯ СПОСОБНОСТЬ КАНАЛА ПРИ ОГРАНИЧЕНИИ СРЕДНЕЙ МОЩНОСТИ

Простым применением теоремы 16 является случай, когда шумы представляют собой «белые» тепловые шумы, а принимаемые сигналы ограничены некоторой средней мощностью P . Тогда принимаемые сигналы имеют среднюю мощность $P + N$, где N есть сред-

няя мощность шумов. Принимаемые сигналы обладают максимальной «энтропией», когда они также образуют ансамбль «белых» шумов, так как это наибольшая возможная «энтропия» для мощности $P+N$. Такая «энтропия» может быть получена путем надлежащего выбора ансамбля передаваемых сигналов, а именно в том случае, если они образуют ансамбль «белых» шумов мощностью P . Тогда «энтропия» (за 1 сек.) принимаемого ансамбля будет

$$H(y) = W \log 2\pi e(P + N),$$

а «энтропия» шумов

$$H(n) = W \log 2\pi eN.$$

Пропускная способность канала равна

$$C = H(y) - H(n) = W \log \frac{P+N}{N}.$$

Теорема 17

Пропускная способность канала с полосой частот W , на которой действуют «белые» тепловые шумы мощностью N , при средней мощности передаваемых сигналов P равна

$$C = W \log \frac{P+N}{N}.$$

Это означает, что при достаточно сложных системах кодирования можно передавать сообщения со скоростью

$$C = W \log_2 \frac{P+N}{N} \frac{\text{двоичных единиц}}{\text{сек.}}$$

при сколь угодно малой частоте ошибок. Невозможно передавать с большей скоростью при любой системе кодирования без того, чтобы частота ошибок не имела конечного положительного значения.

Для достижения этой предельной скорости передаваемые сигналы по своим статистическим свойствам должны приближаться к «белым» шумам. Система, в которой скорости передачи достигают предельной, может быть описана следующим образом. Пусть создаются $M=2^s$ «образцов» «белых» шумов длительностью каждый T . Им приписываются двоичные числа от 0 до $M-1$. В передатчике последовательности сообщений разбиваются на группы по S сообщений и для каждой группы передается как сигнал соответствующий «образец» шумов. При приеме значения M «образцов» известны и действительный принимаемый сигнал, искаженный шумами, сравнивается с каждым из них. «Образец», имеющий наименьшее эффективное отклонение от принимаемого сигнала, выбирается как передаваемый сигнал, после чего восстанавливается соответствующее двоичное число.

Этот процесс эквивалентен выбору наиболее вероятного (в апостериорном смысле) сигнала. Число используемых «образцов» шу-

мов M будет зависеть от допустимой частоты ошибок ϵ , но почти для любого выбора числа «образцов» имеем

$$\lim_{\epsilon \rightarrow 0} \lim_{T \rightarrow \infty} \frac{\log M(\epsilon, T)}{T} = W \log \frac{P+N}{N}.$$

Таким образом, независимо от того, сколь малым выбрано ϵ , можно, выбирая T достаточно большим, приблизиться сколь угодно близко к передаче $TW \log \frac{P+N}{N}$ двоичных единиц за время T .

Формулы, подобные $C = W \log \frac{P+N}{N}$, для случая «белых» шумов были получены независимо и другими авторами, хотя при несколько другой интерпретации.

В случае произвольных мешающих шумов (не обязательно «белых» тепловых шумов) задача разыскания максимума, связанная с определением пропускной способности C , повидимому, не может быть полностью решена. Однако могут быть установлены верхний и нижний пределы для C , выраженные через среднюю мощность шумов N и энтропийную мощность шумов N_1 . В большинстве практических случаев эти пределы достаточно близки друг к другу, и поэтому полученное решение может считаться удовлетворительным.

Теорема 18

Пропускная способность канала с полосой частот W , на который воздействуют произвольные шумы, ограничивается неравенствами

$$W \log \frac{P+N_1}{N_1} \ll C \leq W \log \frac{P+N}{N},$$

где P — средняя мощность передаваемых сигналов;

N — средняя мощность шумов;

N_1 — энтропийная мощность шумов.

Здесь опять средняя мощность искаженных сигналов будет $P+N$. Максимальная «энтропия» была бы в том случае, когда принимаемые сигналы представляли бы собой «белые» шумы и она равнялась бы $W \log 2\pi e(P+N)$. Достигнуть этого невозможно, т. е. не может быть такого ансамбля передаваемых сигналов, который, будучи добавлен к мешающим шумам, создавал бы в приемнике «белые» тепловые шумы. Однако это определяет верхний предел для $H(y)$ и поэтому

$$C = \max H(y) - H(n) < W \log 2\pi e(P+N) - W \log 2\pi e N_1.$$

Это и есть верхний предел, указанный в теореме.

Нижний предел может быть найден при рассмотрении скорости передачи в случае, если передаваемые сигналы представляют собой «белые» шумы мощностью P . При этом энтропийная мощность принимаемых сигналов должна быть, по крайней мере, равна энтропийной

мощности «белых» шумов со средней мощностью $P+N_1$. Действительно, теорема 15 устанавливает, что энтропийная мощность суммы двух ансамблей больше или равна сумме отдельных энтропийных мощностей. Отсюда

$$\max H(y) \geq W \log 2\pi e(P+N_1)$$

и

$$C \geq W \log 2\pi e(P+N_1) - W \log 2\pi eN_1 = W \log \frac{P+N_1}{N_1}.$$

По мере возрастания P верхний и нижний пределы сходятся, поэтому предельная скорость равна

$$W \log \frac{P+N}{N_1}$$

Если сами шумы являются «белыми», то $N=N_1$ и полученный результат сводится к доказанной ранее формуле

$$C = W \log \frac{P+N}{N}.$$

Если шумы подчиняются нормальному закону, но спектр их не обязательно равномерный, то N_1 есть геометрическое среднее мощности шумов, взятое по различным частотам в полосе W . Таким образом,

$$N_1 = \exp \frac{1}{W} \int_W \log N(f) df,$$

где $N(f)$ — мощность шумов на частоте f .

Т е о р е м а 19

Если установить пропускную способность канала при данной мощности передаваемых сигналов P равной

$$C = W \log \frac{P+N-\eta}{N_1},$$

то η монотонно убывает при возрастании P , стремясь в пределе к нулю.

Допустим, что при данной мощности P_1 пропускная способность канала равна

$$W \log \frac{P_1+N-\eta_1}{N_1}.$$

Это означает, что наилучшее возможное распределение сигнала, скажем $p(x)$, будучи добавлено к распределению шумов $q(x)$, дает принимаемое распределение $r(y)$, энтропийная мощность которого есть $P_1+N-\eta_1$. Пусть мощность увеличена до $P_1+\Delta P$ путем добавления к сигналу «белых» шумов мощностью ΔP . «Энтропия» принимаемого сигнала теперь равна по меньшей мере

$$H(y) = W \log 2\pi e(P_1+N-\eta_1+\Delta P),$$

что следует из применения теоремы о минимальной энтропийной мощности суммы. Следовательно, поскольку можно получить указанную величину H , «энтропия» наилучшего распределения должна быть, по крайней мере, такой же, а η должно монотонно убывать. Чтобы показать, что при $P \rightarrow \infty$ $\eta \rightarrow 0$, рассмотрим сигнал, который представляет собой «белые» шумы большой мощности P . Какими бы ни были мешающие шумы, если мощность P достаточно велика, принимаемый сигнал будет приближенно «белыми» шумами в смысле обладания энтропийной мощностью, сходящейся к $P+N$.

25. ПРОПУСКНАЯ СПОСОБНОСТЬ КАНАЛА ПРИ ОГРАНИЧЕНИИ ПИКОВОЙ МОЩНОСТИ

В некоторых случаях ограниченной является не средняя мощность, а мгновенная пиковая. Задача вычисления пропускной способности канала сводится тогда к разысканию максимума (путем вариации ансамбля передаваемых символов) выражения

$$H(y) - H(n)$$

при наложении ограничения, что все функции $f(t)$ в ансамбле для всех t меньше или равны, скажем, \sqrt{S} . Задача при таком ограничении не может быть так же хорошо математически решена, как в случае ограничения средней мощности. В рассматриваемом случае можно определить только нижний предел, пригодный для любых отношений $\frac{S}{N}$, верхний предел, пригодный для больших $\frac{S}{N}$, и приближенное значение C для малых отношений $\frac{S}{N}$.

Теорема 20

Пропускная способность канала C с полосой частот W , на который воздействуют «белые» тепловые шумы мощностью N , ограничена величиной

$$C \geq W \log \frac{2}{\pi e^3} \frac{S}{N},$$

где S — допустимая пиковая мощность передаваемых сигналов. Для достаточно больших отношений $\frac{S}{N}$

$$C \leq W \log \frac{2}{\pi e} \frac{S+N}{N} (1 + \epsilon),$$

где ϵ — сколь угодно мало. При $\frac{S}{N} \rightarrow 0$ и полосе частот W , начинающейся от 0,

$$\frac{C}{W \log \left(1 + \frac{S}{N}\right)} \rightarrow 1.$$

Желательно сделать «энтропию» принимаемых сигналов максимальной. Если $\frac{S}{N}$ велико, то это будет близко соответствовать случаю, когда «энтропия» передаваемого ансамбля максимальна.

Приближенный верхний предел определяется путем ослабления условий, накладываемых на ансамбль. Допустим, что мощность ограничена величиной S не в каждый момент времени, а только в дискретных точках, где фиксируются значения передаваемого сообщения. При таких ослабленных условиях максимальная «энтропия» передаваемого ансамбля будет больше или равна максимальной «энтропии» при исходных условиях. Измененная таким образом задача может быть легко решена. Максимальная «энтропия» имеет место в том случае, если различные «образцы» сигнала независимы и имеют функцию распределения, которая постоянна от $-\sqrt{S}$ до $+\sqrt{S}$ «Энтропия» при этом равна

$$W \log 4S.$$

Принимаемый сигнал будет тогда иметь «энтропию», меньшую, чем

$$W \log(4S + 2\pi eN)(1 + \epsilon),$$

причем $\epsilon \rightarrow 0$ при $\frac{S}{N} \rightarrow \infty$. Пропускную способность канала найдем, вычитая из этого выражения «энтропию» «белых» шумов $W \log 2\pi eN$:

$$W \log(4S + 2\pi eN)(1 + \epsilon) - W \log 2\pi eN = W \log \frac{\frac{2}{\pi e} S + N}{N} (1 + \epsilon).$$

Это и есть искомое выражение для верхнего предела пропускной способности канала.

Чтобы найти нижний предел, рассмотрим тот же самый ансамбль функций. Пусть эти функции проходят через идеальный фильтр с треугольной характеристикой, коэффициент передачи которого равен единице при нулевой частоте и линейно спадает до нуля на частоте W .

Прежде всего покажем, что функции на выходе фильтра ограничены по пиковой мощности во все моменты времени, а не только в указанных дискретных точках. Сначала заметим, что импульс $\frac{\sin 2\pi Wt}{2\pi Wt}$ проходя через фильтр, создает на выходе функцию

$$\frac{1}{2} \frac{\sin^2 \pi Wt}{(\pi Wt)^2},$$

которая никогда не является отрицательной. В общем случае входная функция может рассматриваться как ряд сдвинутых во времени функций

$$a \frac{\sin 2\pi Wt}{2\pi Wt},$$

где амплитуда «образца» a не превышает \sqrt{S} . Следовательно, напряжение на выходе состоит из суммы сдвинутых во времени неотрицательных функций указанного ранее типа с такими же коэффициентами. Для любого момента времени t эти функции принимают наибольшее положительное значение в том случае, когда все коэффициенты a имеют максимальную положительную величину, т. е. \sqrt{S} . В этом случае входная функция есть постоянная с амплитудой \sqrt{S} , а так как фильтр для постоянной составляющей имеет коэффициент передачи единицу, то и выходная функция будет такой же. Таким образом, выходной ансамбль имеет пиковую мощность S .

«Энтропия» выходного ансамбля может быть найдена по «энтропии» входного ансамбля при помощи доказанной ранее теоремы. Выходная «энтропия» равна входной плюс геометрическое среднее коэффициента передачи фильтра

$$\int_0^W \log |Y(f)|^2 df = \int_0^W \log \left(\frac{W-f}{W} \right)^2 df = -2W.$$

Следовательно, выходная «энтропия» равна

$$W \log 4S - 2W = W \log \frac{4S}{e^2},$$

а пропускная способность канала больше, чем

$$W \log \frac{2}{\pi e^3} \frac{S}{N}$$

Теперь надо показать, что для малых значений $\frac{S}{N}$ (отношение пиковой мощности сигнала к средней мощности «белых» шумов) пропускная способность канала равна приблизительно

$$C = W \log \left(1 + \frac{S}{N} \right)$$

Более точно

$$\frac{C}{W \log \left(1 + \frac{S}{N} \right)} \rightarrow 1 \text{ при } \frac{S}{N} \rightarrow 0.$$

Поскольку средняя мощность сигнала P меньше или равна его пиковой мощности S , то отсюда следует, что для всех $\frac{S}{N}$

$$C \leq W \log \left(1 + \frac{P}{N} \right) \leq W \log \left(1 + \frac{S}{N} \right)$$

Поэтому если найден ансамбль функций, соответствующих скорости передачи, близкой к $W \log \left(1 + \frac{S}{N} \right)$, и ограниченных полосой частот

W и пиковой мощностью S , то последняя часть теоремы будет доказана.

Рассмотрим ансамбль функций следующего типа. Последовательность из t «образцов» имеет одно и то же значение — либо $+\sqrt{S}$, либо $-\sqrt{S}$; следующие t «образцов» опять имеют одинаковое значение и т. д. Значения для последовательности выбираются случайно с вероятностью $1/2$ для $+\sqrt{S}$ и $1/2$ для $-\sqrt{S}$. Если этот ансамбль пропустить через фильтр с треугольной характеристикой (коэффициент передачи для постоянной составляющей равен единице), то на выходе пики будут ограничены до S . Кроме того, средняя мощность близка к S и может достигнуть этого значения, если взять t достаточно большим.

«Энтропия» суммы из этого ансамбля и тепловых шумов может быть найдена использованием теоремы о сумме шумов и малого сигнала. Теорема применима, если

$$\sqrt{t} \frac{S}{N}$$

достаточно мало. Это можно обеспечить, взяв отношение $\frac{S}{N}$ достаточно малым (после того как t выбрано). Энтропийная мощность со сколь угодно близким приближением будет равна $S+N$, и, следовательно, скорость передачи со сколь угодно близка к

$$W \log \left(\frac{S+N}{N} \right).$$

Глава V

СКОРОСТЬ СОЗДАНИЯ СООБЩЕНИЙ ДЛЯ НЕПРЕРЫВНОГО ИСТОЧНИКА

26. ФУНКЦИИ ОЦЕНКИ ВЕРНОСТИ ВОСПРОИЗВЕДЕНИЯ

В случае дискретного источника была определена конечная скорость создания сообщений, а именно «энтропия» соответствующего стохастического процесса. Для непрерывного источника положение оказывается значительно более сложным. Прежде всего непрерывно изменяющаяся величина предполагает бесконечное число значений и поэтому для точного задания требует бесконечного числа двоичных знаков. Это означает, что при передаче выходного эффекта непрерывного источника для *точного воспроизведения* сообщения в месте приема, вообще говоря, необходим канал с бесконечной пропускной способностью. Поскольку в каналах существует обычно определенный уровень шумов и, следовательно, пропускная способность ограничена, точная передача невозможна.

Это рассуждение, однако, обходит действительное положение вещей. Практически при непрерывном источнике может интересоваться не точная передача, а передача с определенным допуском. Вопрос заключается в том, можно ли приписать непрерывному источнику конечную скорость в том случае, когда требуется только определенная верность воспроизведения, измеренная подходящим способом. Разумеется, при возрастании требований к верности воспроизведения скорость создания сообщений будет возрастать.

Как будет показано, в весьма общих случаях можно определить такую скорость. Путем надлежащего кодирования создаваемые сообщения можно передать по каналу, пропускная способность которого равна рассматриваемой скорости, и выполнить при этом требования к верности воспроизведения. Канал, обладающий меньшей пропускной способностью, такой возможности не обеспечивает.

Прежде всего необходимо дать общую математическую формулировку понятию о верности передачи. Рассмотрим группу сообщений большой длительности, скажем T секунд. Источник описывается заданием в соответствующем пространстве плотности вероятностей $P(x)$ того, что будет выбрано рассматриваемое сообщение. Данная система связи описывается (с внешней точки зрения) заданием условной вероятности $P_x(y)$ того, что если источник создал сообщение x , воспроизводимое сообщение в месте приема будет y . Система в целом (включая источник и передающую систему) опи-

сывается функцией вероятностей $P(x, y)$ наличия передаваемого сообщения x и принимаемого сообщения y . Если эта функция известна, то тем самым полностью известны свойства системы с точки зрения верности воспроизведения.

Любая оценка верности должна математически соответствовать операции над функцией $P(x, y)$. Эта операция должна, по крайней мере, давать сравнительную оценку системы. Другими словами, необходимо, чтобы в результате можно было сказать, что согласно нашему критерию верности из двух систем, описываемых функциями $P_1(x, y)$ и $P_2(x, y)$, либо: 1) первая обеспечивает более высокую верность; 2) вторая обеспечивает более высокую верность, либо 3) они обеспечивают одинаковую верность. Это значит, что критерий верности может быть представлен численно оцениваемой функцией

$$v[P(x, y)],$$

аргумент которой изменяется по возможным функциям вероятностей $P(x, y)$. В дальнейшем будем полагать, что меньшим значениям функции оценки соответствует более высокая верность.

Теперь покажем, что при очень общих и приемлемых допущениях функция $v[P(x, y)]$ может быть написана в значительно более специализированной форме, а именно как среднее функции $\rho(x, y)$, взятое по множеству возможных значений x и y :

$$v[P(x, y)] = \iint P(x, y) \rho(x, y) dx dy.$$

Чтобы это показать, достаточно предположить: 1) что источник и система являются эргодическими, так что очень длительный «образец» сообщения будет с вероятностью, близкой к единице, типичен для ансамбля, и 2) что оценка является «приемлемой» в том смысле, что возможно на основе наблюдения типичных входных и выходных «образцов» x_1 и y_1 создать опытную оценку и, если длительность этих «образцов» возрастает, опытная оценка будет с вероятностью единица сходиться к точной оценке, основанной на полном знании функции $P(x, y)$.

Пусть опытная оценка будет $\rho(x, y)$. Тогда функция $\rho(x, y)$ при $T \rightarrow \infty$ стремится к постоянной величине почти для всех значений (x, y) , которые находятся в области высокой вероятности для данной системы

$$\rho(x, y) \rightarrow v[P(x, y)],$$

и можно также написать

$$\rho(x, y) \rightarrow \iint P(x, y) \rho(x, y) dx dy,$$

так как

$$\iint P(x, y) dx dy = 1.$$

Это доказывает искомый результат.

Функция $\rho(x, y)$ имеет общую природу «расстояния» между x и y^1). Она измеряет, насколько было бы плохо (относительно нашего критерия верности) принять y , когда передано x . Полученный выше общий результат может быть сформулирован еще следующим образом. Любая приемлемая оценка может быть представлена как среднее функции «расстояния», взятое по множеству исходных и воспроизводимых сообщений x и y и взвешенное в соответствии с вероятностью $P(x, y)$ получения рассматриваемых пар, полагая, что длительность сообщений T взята достаточно большой.

Ниже даются простые примеры функций оценки.

1. Эффективный критерий

$$v = \overline{[x(t) - y(t)]^2}.$$

В этом очень часто применяемом критерии верности функция «расстояния» $\rho(x, y)$ представляет собой (отвлекаясь от постоянного множителя) квадрат обычного евклидова расстояния между точками x и y в соответствующем функциональном пространстве

$$\rho(x, y) = \frac{1}{T} \int_0^T [x(t) - y(t)]^2 dt.$$

2. Частотно-взвешенный эффективный критерий. Прежде чем воспользоваться эффективной мерой верности, можно приписать различным частотным составляющим разные веса. Это эквивалентно пропусканию разности $x(t) - y(t)$ через формирующий фильтр с последующим определением средней мощности на выходе. Положим

$$e(t) = x(t) - y(t)$$

и

$$f(t) = \int_{-\infty}^{\infty} e(\tau) k(t - \tau) d\tau,$$

тогда

$$\rho(x, y) = \frac{1}{T} \int_0^T f(t)^2 dt.$$

3. Критерий абсолютной ошибки

$$\rho(x, y) = \frac{1}{T} \int_0^T |x(t) - y(t)| dt.$$

¹⁾ Однако она не является «метрической» в строгом смысле, так как в общем случае не удовлетворяет либо условию $\rho(x, y) = \rho(y, x)$, либо условию $\rho(x, y) + \rho(y, z) \geq \rho(x, z)$.

4. Свойства слуха определяют оценку или, скорее, ряд оценок, подходящих в случае передачи речи или музыки. Так, например, существует критерий разборчивости, при котором $\rho(x, y)$ равна относительной частоте неправильно интерпретированных слов, когда сообщение $x(t)$ принимается как $y(\hat{t})$. Хотя в этих случаях нельзя дать точного представления функции $\rho(x, y)$, она может быть в принципе определена путем эксперимента. Некоторые ее свойства следуют из хорошо известных экспериментальных результатов по исследованию слуха, например из того, что ухо сравнительно нечувствительно к фазе, а чувствительность к амплитуде и частоте приближенно логарифмическая.

5. Дискретный случай может рассматриваться как частный случай, в котором подразумевается оценка, основанная на частоте ошибок. Функция $\rho(x, y)$ определяется тогда как число символов в последовательности y , отличающихся от соответствующих символов в последовательности x , деленное на полное число символов в последовательности x .

27. СКОРОСТЬ ИСТОЧНИКА ПРИ ДАННОЙ ОЦЕНКЕ ВЕРНОСТИ

Теперь можно определить скорость создания сообщений для непрерывного источника. Дана функция $P(x)$ для источника и оценка v , определяемая функцией «расстояния» $\rho(x, y)$, которая предполагается непрерывной по x и y . Для данной системы $P(x, y)$ качество измеряется величиной

$$v = \iint \rho(x, y) P(x, y) dx dy.$$

Кроме того, скорость выдачи двоичных единиц, соответствующая данной функции $P(x, y)$, равна

$$R = \iint P(x, y) \log \frac{P(x, y)}{P(x)P(y)} dx dy$$

Определим скорость создания сообщений R_1 при данном качестве воспроизведения v_1 как минимум R , полученный при фиксированном $v = v_1$ и при изменении $P_x(y)$. Таким образом,

$$R_1 = \min_{P_x(y)} \iint P(x, y) \log \frac{P(x, y)}{P(x)P(y)} dx dy$$

при условии

$$v_1 = \iint P(x, y) \rho(x, y) dx dy.$$

Это означает, что в действительности рассматриваются все системы связи, которые могли бы быть использованы и которые обеспечивают передачу с требуемой верностью. Скорость передачи в $\frac{\text{двоичных единицах}}{\text{сек.}}$ вычисляется для каждой системы и выбирается

ся наименьшая скорость. Эта последняя и есть скорость, приписываемая источнику при рассматриваемой верности воспроизведения.

Обоснование этого определения заключается в следующей теореме.

Теорема 21

Если источник при данной оценке v_1 имеет скорость создания сообщений R_1 , то можно закодировать сообщения на выходе источника и передавать их по каналу с пропускной способностью C при верности воспроизведения, как угодно близкой к v_1 , если только $R_1 \leq C$. Это невозможно, если $R_1 > C$.

Последнее утверждение теоремы немедленно следует из определения R_1 и предыдущих результатов. Если оно не справедливо, то можно было бы передавать больше чем C $\frac{\text{двоичных единиц}}{\text{сек.}}$

по каналу с пропускной способностью C .

Первая часть теоремы доказывается методом, аналогичным использованному при доказательстве теоремы 11. Прежде всего следует разделить пространство (x, y) на большое число малых ячеек и рассматривать этот случай как дискретный. Это изменит функцию оценки не больше чем на произвольно малую величину (если ячейки весьма малы) вследствие предполагаемой для функции $\rho(x, y)$ непрерывности. Допустим, что $P_1(x, y)$ есть частная система, при которой скорость минимальна и равна R_1 . Выберем из высоковероятных сообщений y по произволу некоторый ряд, содержащий

$$2^{(R_1 + \epsilon)T}$$

членов, где при $T \rightarrow \infty$ $\epsilon \rightarrow 0$. При большом T каждая выбранная точка будет соединена линией высокой вероятности (как на фиг. 9) с рядом x . Вычисления, подобные использованным при доказательстве теоремы 11, показывают, что при большом T почти все x охватываются «веерами» линий, идущими от выбранных точек y , почти при любом выборе y .

Соответствующая система связи действует следующим образом. Выбранным точкам приписываются двоичные числа. Когда появляется сообщение x , оно будет (с вероятностью, достигающей 1 при $T \rightarrow \infty$) расположено, по крайней мере, на одном из «вееров» линий. Тогда по каналу передается соответствующее двоичное число (или, если их несколько, одно произвольно выбранное число), закодированное надлежащим образом для обеспечения малой вероятности ошибок. Это возможно, поскольку $R_1 \leq C$. В приемной точке восстанавливается соответствующее y , которое и используется как принимаемое сообщение.

Оценка v'_1 для этой системы может быть сделана сколь угодно близкой к v_1 , если взять T достаточно большим. Это обусловливается

тем, что для каждого длинного «образца» сообщения $x(t)$ и воспроизводимого сообщения $y(t)$ оценка сходится к v_1 (с вероятностью 1).

Интересно отметить, что в этой системе шумы в воспроизводимом сообщении создаются за счет специфического квантования в передатчике, а не за счет шумов в канале. Они в некоторой степени аналогичны шумам квантования при кодовой импульсной модуляции.

28. ВЫЧИСЛЕНИЕ СКОРОСТИ СОЗДАНИЯ СООБЩЕНИЙ

Определение скорости создания сообщений во многих отношениях подобно определению пропускной способности канала. В первом случае

$$R = \min_{P_x(y)} \iint P(x, y) \log \frac{P(x, y)}{P(x)P(y)} dx dy$$

при фиксированных $P(x)$ и $v_1 = \iint P(x, y) \rho(x, y) dx dy$.

Во втором случае

$$C = \max_{P(x)} \iint P(x, y) \log \frac{P(x, y)}{P(x)P(y)} dx dy$$

при фиксированном $P_x(y)$ и при наложении, возможно, одного или более других ограничений (например, ограничение средней мощности) вида $K = \iint P(x, y) \lambda(x, y) dx dy$.

Для общей задачи разыскания максимума, возникающей при определении скорости источника, может быть найдено частное решение. Пользуясь методом Лагранжа, рассмотрим выражение

$$\iint \left[P(x, y) \log \frac{P(x, y)}{P(x)P(y)} + \mu P(x, y) \rho(x, y) + \nu(x) P(x, y) \right] dx dy.$$

Вариационное уравнение, когда берут первую вариацию по $P(x, y)$, приводит к

$$P_y(x) = B(x) e^{-\lambda \rho(x, y)},$$

где λ определяется из условия получения необходимой верности воспроизведения, а $B(x)$ должно удовлетворять равенству

$$\int B(x) e^{-\lambda \rho(x, y)} dx = 1.$$

Это показывает, что при наилучшем кодировании условная вероятность для различных принимаемых сообщений y , т. е. $P_y(x)$, экспоненциально уменьшается вместе с $\rho(x, y)$ — функцией «расстояния» между рассматриваемыми x и y .

В частном случае, когда функция «расстояния» $\rho(x, y)$ зависит только от (векторной) разности между x и y

$$\rho(x, y) = \rho(x - y),$$

имеем

$$\int B(x) e^{-\lambda \rho(x-y)} dx = 1.$$

Следовательно, $B(x)$ — постоянная величина, скажем α , и

$$P_y(x) = \alpha e^{-\lambda \rho(x-y)}.$$

К сожалению, эти формальные решения в частных случаях трудно численно оценить, и поэтому ценность их представляется небольшой. Фактическое вычисление скоростей источников было выполнено только для немногих очень простых случаев.

Если функция «расстояния» $\rho(x, y)$ представляет собой средний квадрат разности между x и y , а ансамбль сообщений — «белые» шумы, то скорость может быть определена. В этом случае имеем

$$R = \min [H(x) - H_y(x)] = H(x) - \max H_y(x)$$

при $N = \overline{(x-y)^2}$. Но максимум $H_y(x)$ соответствует случаю, когда $y-x$ есть «белые» шумы, причем он равен $W_1 \log 2\pi eN$, где W_1 — полоса частот ансамбля сообщений. Поэтому

$$R = W_1 \log 2\pi eQ - W_1 \log 2\pi eN = W_1 \log \frac{Q}{N},$$

где Q — средняя мощность сообщений. Это доказывает следующую теорему.

Теорема 22

Скорость источника «белых» шумов мощностью Q и с полосой частот W_1 при эффективном критерии верности воспроизведения равна

$$R = W_1 \log \frac{Q}{N},$$

где N есть допустимый средний квадрат отклонения воспроизводимого сообщения от исходного сообщения.

В более общем случае для любого источника сообщений можно получить неравенства, ограничивающие скорость создания сообщений при допустимом среднем квадрате отклонения.

Теорема 23

Скорость для любого источника с полосой частот W_1 ограничена соотношениями

$$W_1 \log \frac{Q_1}{N} \leq R \leq W_1 \log \frac{Q}{N},$$

где Q есть средняя мощность источника, Q_1 — его энтропийная мощность и N — допустимый средний квадрат отклонения.

Нижний предел следует из того, что максимум $H_y(x)$ при данном $\overline{(x-y)^2} = N$ имеет место в случае «белых» шумов. Верхний предел будет получен, если разместить точки, использованные при доказательстве теоремы 21, не лучшим образом, а случайно в сфере радиуса $\sqrt{Q-N}$

Приложение 1

Пусть $N_i(L)$ будет число групп символов длительностью L , заканчивающихся в состоянии i . Тогда имеем

$$N_j(L) = \sum_{i,s} N_i \left(L - b_{ij}^{(s)} \right),$$

где $b_{ij}^1, b_{ij}^2, \dots, b_{ij}^m$ — длительности символов, которые могут быть выбраны в состоянии i и приводят к состоянию j . Эти выражения суть линейные разностные уравнения, свойства которых при $L \rightarrow \infty$ должны быть вида

$$N_j = A_j W^L$$

Подставляем в разностное уравнение

$$A_j W^L = \sum_{i,s} A_i W^{L - b_{ij}^{(s)}}$$

или

$$A_j = \sum_{i,s} A_i W^{-b_{ij}^{(s)}}$$

$$\sum_i \left(\sum_s W^{-b_{ij}^{(s)}} - \delta_{ij} \right) A_i = 0.$$

Чтобы это было возможно, определитель

$$D(W) = |a_{ij}| = \left| \sum_s W^{-b_{ij}^{(s)}} - \delta_{ij} \right|$$

должен быть равен нулю, что дает W , который, конечно, является наибольшим действительным корнем уравнения $D=0$.

Тогда величина C равна

$$C = \lim_{L \rightarrow \infty} \frac{\log \sum A_j W^L}{L} = \log W$$

Замечаем, что придем к тому же результату, если потребуем, чтобы все группы начинались с одного и того же произвольно выбранного состояния.

Приложение 2

Пусть $H\left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\right) = A(n)$. По условию (3) можно разбить выбор из числа s^m равных возможностей на ряды по m выборов из s равных возможностей в каждом и получить

$$A(s^m) = mA(s).$$

Точно так же

$$A(t^n) = nA(t).$$

Можно выбрать n произвольно большим и найти m из условия

$$s^m \leq t^n < s^{m+1}.$$

Таким образом, логарифмируя и деля на $n \log s$, найдем

$$\frac{m}{n} \leq \frac{\log t}{\log s} < \frac{m}{n} + \frac{1}{n}$$

или

$$\left| \frac{m}{n} - \frac{\log t}{\log s} \right| < \epsilon,$$

где ϵ — произвольно мало.

Теперь из свойств монотонности $A(n)$

$$A(s^m) \leq A(t^n) \leq A(s^{m+1}),$$

$$mA(s) \leq nA(t) \leq (m+1)A(s).$$

Следовательно, деля на $nA(s)$:

$$\frac{m}{n} \leq \frac{A(t)}{A(s)} \leq \frac{m}{n} + \frac{1}{n}$$

или

$$\left| \frac{m}{n} - \frac{A(t)}{A(s)} \right| < \epsilon,$$

$$\left| \frac{A(t)}{A(s)} - \frac{\log t}{\log s} \right| \leq 2\epsilon, \quad A(t) = -K \log t,$$

где K должно быть положительным, чтобы удовлетворить условию (2).

Допустим теперь, что имеется выбор из n возможностей с измеримыми вероятностями $p_i = \frac{n_i}{\sum n_i}$, где n_i — целые числа.

Можно разделить выбор из $\sum n_i$ возможностей на выбор из n возможностей с вероятностями p_1, \dots, p_n и затем, если i было избрано, произвести выбор из n_i возможностей с равными вероятностями. Пользуясь опять условием (3), приравняем полный выбор из $\sum n_i$ возможностей, вычисленный двумя способами

$$K \log \sum n_i = H(p_1, \dots, p_n) + K \sum p_i \log n_i.$$

Следовательно:

$$H = K \left(\sum p_i \log \sum n_i - \sum p_i \log n_i \right) =$$

$$= -K \sum p_i \log \frac{n_i}{\sum n_i} = -K \sum p_i \log p_i.$$

Если p_i иррациональны, они могут быть аппроксимированы правильными дробями и то же самое выражение должно сохраниться при предположении о непрерывности. Таким образом, это выражение справедливо в общем случае. Выбор коэффициента K производится из соображений удобства, он определяет единицу измерений.

Приложение 3

Предположим, что источник является эргодическим, так что применим сильный закон больших чисел. Таким образом, число пересечений данной траектории p_{ij} в последовательности большой длины N приблизительно пропорционально вероятности нахождения в состоянии i (скажем, P_i) и последующего выбора этой траектории, т. е. $P_i p_{ij} N$. Если N достаточно велико, то вероятность ошибки $\pm \delta$ при этом меньше ϵ , так что для всех случаев за исключением группы малой вероятности действительные числа заключены в пределах

$$(P_i p_{ij} \pm \delta) N.$$

Следовательно, почти все последовательности имеют вероятность

$$p = \prod p_{ij}^{(P_i p_{ij} \pm \delta) N}$$

и $\frac{\log p}{N}$ ограничен величиной

$$\frac{\log p}{N} = \sum (P_i p_{ij} \pm \delta) \log p_{ij},$$

или

$$\left| \frac{\log p}{N} - \sum P_i p_{ij} \log p_{ij} \right| < \eta.$$

Это доказывает теорему 3.

Теорема 4 немедленно следует отсюда по вычислении верхнего и нижнего пределов для $n(q)$, основанных на диапазоне возможных значений p в теореме 3.

В смешанном (не эргодическом) случае, если

$$L = \sum p_i L_i,$$

а «энтропии» составляющих суть $H_1 \geq H_2 \geq \dots \geq H_n$, справедливо следующее предложение.

Теорема

$$\lim_{N \rightarrow \infty} \frac{\log n(q)}{N} = \varphi(q)$$

есть убывающая ступенчатая функция

$$\varphi(q) = H_s \text{ в интервале } \sum_1^{s-1} a_i < q < \sum_1^s a_i.$$

Для доказательства теорем 5 и 6 прежде всего заметим, что F_N монотонно убывает, так как увеличение N увеличивает индекс условной «энтропии». Простая подстановка значения $p_{B_i}(S_j)$ в формулу для F_N показывает, что

$$F_N = NG_N - (N-1)G_{N-1}.$$

Суммируя по всем N , получим

$$G_N = \frac{1}{N} \sum F_N.$$

Следовательно, $G_N \geq F_N$ и G_N монотонно убывает. Они должны также сходиться к одному и тому же пределу. Пользуясь теоремой 3, видим, что

$$\lim_{N \rightarrow \infty} G_N = H.$$

Приложение 4

Допустим, что имеется ряд ограничений, наложенных на последовательности символов, причем последовательности — с конечными состояниями и поэтому могут быть представлены линейным графиком, как на фиг 2. Пусть $l_{ij}^{(s)}$ будут длительности различных символов, которые могут случиться при переходе из состояния i в состояние j . Какое распределение вероятностей P_i для различных состояний и вероятностей $p_{ij}^{(s)}$ выбора символа s в состоянии i , переходящем в состояние j , дает максимальную скорость создания сообщений при данных ограничениях? Ограничения определяют дискретный канал, а максимальная скорость должна быть меньше или равна пропускной способности C этого канала. Действительно, если все группы большой длительности равновероятны, то в результате получилась бы именно эта скорость, а если они возможны, то такая скорость была бы наилучшей. Ниже будет показано, что эта скорость может быть получена путем надлежащего выбора P_i и $p_{ij}^{(s)}$

Рассматриваемая скорость равна

$$\frac{- \sum_{i,j,s} P_i p_{ij}^{(s)} \log p_{ij}^{(s)}}{\sum_{i,j,s} P_i p_{ij}^{(s)} l_{ij}^{(s)}}$$

Пусть

$$p_{ij}^{(s)} = \frac{B_j}{B_i} W^{-l_{ij}^{(s)}}$$

где B_i удовлетворяет уравнению

$$B_i = \sum_{j,s} B_j W^{-l_{ij}^{(s)}}$$

Эта однородная система имеет не равное нулю решение, поскольку W таково, что определитель коэффициентов равен нулю

$$\left| \sum_s W^{-l_{ij}^{(s)}} - \delta_{ij} \right| = 0.$$

Выбранные таким образом $p_{ij}^{(s)}$ являются подходящими переходными вероятностями, так как прежде всего

$$\sum_{j,s} p_{ij}^{(s)} = \sum_{j,s} \frac{B_j}{B_i} W^{-l_{ij}^{(s)}} = \frac{B_i}{B_i} = 1,$$

так что сумма вероятностей в любой частной узловой точке равна единице. Далее, они не отрицательны, как это можно видеть из рассмотрения величин A_i (Приложение 1). Все A_i обязательно не отрицательны, а B_i удовлетворяют подобной же системе уравнений, только с переменной местами i и j . Это приводит к обратной ориентации линий на графике.

Подставляя эти значения $p_{ij}^{(s)}$ в общее уравнение для скорости, получим

$$\begin{aligned} & \frac{\sum P_i p_{ij}^{(s)} \log \frac{B_j}{B_i} W^{-l_{ij}^{(s)}}}{\sum P_i p_{ij}^{(s)} l_{ij}^{(s)}} = \\ & = \frac{\log W \sum P_i p_{ij}^{(s)} l_{ij}^{(s)} - \sum P_i p_{ij}^{(s)} \log B_j + \sum P_i p_{ij}^{(s)} \log B_i}{\sum P_i p_{ij}^{(s)} l_{ij}^{(s)}} = \log W = C. \end{aligned}$$

Таким образом, скорость при этой группе переходных вероятностей равна C , и поскольку эта скорость никогда не может быть превзойдена, то она является максимальной.

Приложение 5

Пусть S_1 будет некоторое измеримое подмножество g -ансамбля, а S_2 — подмножество f -ансамбля, которое дает S_1 в результате операции T . Тогда

$$S_1 = TS_2.$$

Пусть H^λ будет оператор, смещающий на интервал времени λ все функции множества. Тогда

$$H^\lambda S_1 = H^\lambda TS_2 = TH^\lambda S_2,$$

так как T — инвариантна и поэтому может переставляться с H^λ . Таким образом, если $m[S]$ есть вероятностная мера множества S , то

$$m[H^\lambda S_1] = m[TH^\lambda S_2] = m[H^\lambda S_2] = m[S_2] = m[S_1],$$

где второе равенство следует из определения меры в g -пространстве, третье — из стационарности f -ансамбля, а последнее — опять-таки из определения меры g . Это показывает, что g -ансамбль — стационарный.

Для доказательства сохранения эргодических свойств при инвариантных операциях положим, что S_1 есть подмножество g -ансамбля, инвариантное при операциях H^λ , и пусть S_2 будет множество всех функций f , которые преобразуются в S_1 . Тогда

$$H^\lambda S_1 = H^\lambda T S_2 = T H^\lambda S_2 = S_1,$$

так что $H^\lambda S_2$ включается в S_2 при всех λ . Теперь так как

$$m[H^\lambda S_2] = m[S_2] = m[S_1],$$

то это означает

$$H^\lambda S_2 = S_2$$

для всех λ при $m[S_2] \neq 0, 1$. Это противоречие показывает, что S_1 не существует.

Приложение 6

Верхний предел $\bar{N}_3 \leq N_1 + N_2$ объясняется тем обстоятельством, что максимальная возможная «энтропия» для мощности $N_1 + N_2$ будет в том случае, когда имеются «белые» шумы такой мощности. При этом энтропийная мощность равна $N_1 + N_2$.

Чтобы найти нижний предел, допустим, что имеются два распределения в n измерениях $p(x_i)$ и $q(x_i)$ с энтропийными мощностями \bar{N}_1 и \bar{N}_2 . Какую форму должны иметь p и q , чтобы энтропийная мощность \bar{N}_3 их взаимодействия

$$r(x_i) = \int p(y_i) q(x_i - y_i) dy_i$$

была минимальной?

«Энтропия» для r , которую обозначим H_3 , равна

$$H_3 = - \int r(x_i) \lg r(x_i) dx_i.$$

Надо разыскать минимум этого выражения при наложении следующих условий:

$$H_1 = - \int p(x_i) \lg p(x_i) dx_i,$$

$$H_2 = - \int q(x_i) \lg q(x_i) dx_i.$$

Рассмотрим тогда

$$U = - \int [r(x) \log r(x) + \lambda p(x) \log p(x) + \mu q(x) \log q(x)] dx,$$

$$\delta U = - \int \{ [1 + \log r(x)] \delta r(x) + \lambda [1 + \log p(x)] \delta p(x) + \\ + \mu [1 + \log q(x)] \delta q(x) \} dx.$$

Если $p(x)$ изменяется при частном значении аргумента $x_i = s_i$, изменение $r(x)$ равно

$$\delta r(x) = q(x_i - s_i)$$

и

$$\delta U = - \int q(x_i - s_i) \log r(x_i) dx_i - \lambda \log p(s_i) = 0.$$

Так же точно обстоит дело, если изменяется q . Таким образом, условия минимума

$$\int q(x_i - s_i) \log r(x_i) dx_i = - \lambda \log p(s_i),$$

$$\int p(x_i - s_i) \log r(x_i) dx_i = - \mu \log q(s_i).$$

Если умножить первое выражение на $p(s_i)$, а второе на $q(s_i)$ и проинтегрировать по s , то можно получить

$$H_3 = - \lambda H_1,$$

$$H_3 = - \mu H_2$$

или, решая относительно λ и μ и подставляя в уравнения:

$$H_1 \int q(x_i - s_i) \log r(x_i) dx_i = - H_3 \log p(s_i),$$

$$H_2 \int p(x_i - s_i) \log r(x_i) dx_i = - H_3 \log q(s_i).$$

Допустим теперь, что $p(x_i)$ и $q(x_i)$ подчиняются нормальному закону

$$p(x_i) = \frac{|A_{ij}|^{n/2}}{(2\pi)^{n/2}} \exp \left(- \frac{1}{2} \sum A_{ij} x_i x_j \right),$$

$$q(x_i) = \frac{|B_{ij}|^{n/2}}{(2\pi)^{n/2}} \exp \left(- \frac{1}{2} \sum B_{ij} x_i x_j \right).$$

Тогда $r(x_i)$ будет также подчиняться нормальному закону с квадратичной формой C_{ij} . Если обратные величины этих форм составляют a_{ij} , b_{ij} , c_{ij} , то

$$c_{ij} = a_{ij} + b_{ij}.$$

Надо показать, что эти функции удовлетворяют условиям минимума только в том случае, если $a_{ij} = K b_{ij}$, и таким образом дают минимум H_3 при наложенных ограничениях.

Прежде всего имеем

$$\begin{aligned} \log r(x_i) &= \frac{n}{2} \log \frac{1}{2\pi} |C_{ij}| - \frac{1}{2} \sum C_{ij} x_i x_j \\ &\int q(x_i - s_i) \log r(x_i) dx_i = \\ &= \frac{n}{2} \log \frac{1}{2\pi} |C_{ij}| - \frac{1}{2} \sum C_{ij} s_i s_j - \frac{1}{2} \sum C_{ij} b_{ij}. \end{aligned}$$

Это должно равнятьсяся

$$\frac{H_3}{H_1} \left(\frac{n}{2} \log \frac{1}{2\pi} |A_{ij}| - \frac{1}{2} \sum A_{ij} s_i s_j \right),$$

что требует

$$A_{ij} = \frac{H_1}{H_3} C_{ij}.$$

В этом случае $A_{ij} = \frac{H_1}{H_2} B_{ij}$ и оба уравнения превращаются в тождества.

ТЕОРИЯ ФЛУКТУАЦИОННЫХ ШУМОВ¹⁾

С. РАЙС

Глава I

ДРОБОВОЙ ЭФФЕКТ

Дробовой эффект в электронных лампах представляет собой типичный пример шумов. Эти шумы являются следствием флуктуаций интенсивности потока электронов, текущего от катода к аноду. Здесь будет рассмотрен упрощенный тип дробового эффекта.

1.1. ВЕРОЯТНОСТЬ ПОСТУПЛЕНИЯ НА АНОД ТОЧНО K ЭЛЕКТРОНОВ ЗА ПРОМЕЖУТОК ВРЕМЕНИ T

Предположим, что флуктуации электронного потока являются беспорядочными, и будем трактовать эту случайность следующим образом. Подсчитаем число электронов, поступающих на анод за длительный промежуток времени T , измеряемый в секундах. Допустим, что оно равно K_1 . Повторяя процесс подсчета для многих промежутков длительностью T , получим ряд чисел K_2, K_3, \dots, K_M , где M — полное число таких промежутков. Среднее число электронов в секунду ν определяется как

$$\nu = \lim_{M \rightarrow \infty} \frac{K_1 + K_2 + \dots + K_M}{MT}, \quad (1.1-1)$$

причем предполагаем, что этот предел существует. По мере увеличения M при постоянной величине T некоторые K будут иметь одинаковые значения. Действительно, при возрастании M число K , имеющих какое-то данное значение, будет стремиться к увеличению. Это заключение основано на предположении, что электронный ток представляет собой постоянный ток, на который накладываются беспорядочные флуктуации. Вероятность попадания на анод K электронов за данный опыт определяется как

$$p(K) = \lim_{M \rightarrow \infty} \frac{\text{число опытов, дающих точно } K \text{ электронов}}{M} \quad (1.1-2)$$

¹⁾ S. O. Rice, «Mathematical Analysis of Random Noise», *Bell System Technical Journal*, 23, № 3, 282—332, July 1944, 24, № 1, 46—156, January 1945.

Конечно, $p(K)$ зависит также от T . Предположим, что беспорядочность электронного потока такова, что вероятность попадания электрона на анод в промежуток времени $(t, t + \Delta t)$ равна $\nu \Delta t$ (Δt таково, что $\nu \Delta t \ll 1$) и что эта вероятность не зависит от событий, происходящих до момента времени t или после момента $t + \Delta t$.

Этого предположения достаточно для написания выражения для $p(K)$, которое равно

$$p(K) = \frac{(\nu T)^K}{K!} e^{-\nu T}. \quad (1.1-3)$$

Это есть «закон малых вероятностей» Пуассона. Один из иногда применяемых методов доказательства может быть легко продемонстрирован для случая $K=0$. Разделим промежуток $(0, T)$ на M интервалов каждый длительностью $\Delta t = \frac{T}{M}$. Выбираем Δt таким малым, чтобы $\nu \Delta t$ было значительно меньше единицы. (Это и есть «малая вероятность» того, что электрон попадает на анод в промежуток времени Δt .) Вероятность того, что электрон не попадает на анод в первый промежуток Δt , равна $(1 - \nu \Delta t)$. Соответственно вероятность того, что электрон не попадает на анод ни в первый, ни во второй промежуток, равна $(1 - \nu \Delta t)^2$, а вероятность того, что электрон не попадает на анод ни в одном из M промежутков, равна $(1 - \nu \Delta t)^M$. Заменяя M через $\frac{T}{\Delta t}$ и полагая $\Delta t \rightarrow 0$, получим

$$p(0) = e^{-\nu T}$$

Выражения для $p(1), p(2), \dots, p(K)$ могут быть выведены подобным же образом.

1.2. ТЕОРЕМА О НАЛОЖЕНИИ СЛУЧАЙНЫХ ВОЗМУЩЕНИЙ

Допустим, что попадание электрона на анод в момент $t=0$ вызывает какой-то эффект $F(t)$ в некоторой точке выходной цепи. Если выходной контур таков, что эффекты, вызываемые отдельными электронами, складываются линейно, то полный эффект в момент t благодаря действию всех электронов равен

$$I(t) = \sum_{k=-\infty}^{+\infty} F(t - t_k), \quad (1.2-1)$$

где k -й электрон попадает в момент t_k , а ряд предполагается сходящимся.

Теорема о наложении случайных возмущений утверждает, что среднее значение $I(t)$ равно

$$\overline{I(t)} = \nu \int_{-\infty}^{+\infty} F(t) dt, \quad (1.2-2)$$

а средний квадрат флуктуаций вокруг этого среднего значения

$$\overline{[I(t) - \bar{I}(t)]^2} = \nu \int_{-\infty}^{+\infty} F^2(t) dt, \quad (1.2-3)$$

где ν есть среднее число электронов, попадающих на анод за 1 сек.

Формулировка этой теоремы не будет достаточна, пока не определено, что подразумевается под словами «среднее значение». Форма приведенных выше уравнений может натолкнуть на мысль, что это есть среднее во времени, т. е. значение

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T I(t) dt. \quad (1.2-4)$$

Однако при доказательстве этой теоремы усреднение обычно производится по весьма большому числу промежутков длительностью T , а t сохраняется постоянным. Этот процесс в известной степени подобен применявшемуся в разделе (1.1); для большей ясности рассмотрим, что представляет собой, например, $\bar{I}(t)$. Мы наблюдаем $I(t)$ во многих, скажем M , интервалах, длительностью каждый T , причем T велико по сравнению с промежутком, в течение которого эффект $F(t)$, вызываемый поступлением одиночного электрона, значителен. Пусть ${}_n I(t')$ будет значение $I(t)$ через t' секунд после начала n -го интервала. t' равно t плюс постоянная величина, зависящая от начального момента интервала. Индекс поставлен впереди, чтобы сохранить обычное место для другого индекса, который будет введен позже. Значение $\bar{I}(t')$ тогда определяется так:

$$\bar{I}(t') = \lim_{M \rightarrow \infty} \frac{1}{M} [{}_1 I(t') + {}_2 I(t') + \dots + {}_M I(t')], \quad (1.2-5)$$

причем предполагается, что предел существует. Средний квадрат флуктуаций $I(t')$ определяется таким же точно путем.

Как показывают уравнения (1.2-2) и (1.2-3), эти средние значения и все им подобные средние значения, появляющиеся в дальнейшем, оказываются независимыми от времени. Когда это справедливо и входящие в уравнение (1.2-5) M интервалов следуют друг за другом, среднее во времени (1.2-4) и среднее (1.2-5) становятся одинаковыми. Чтобы это показать, умножим обе части (1.2-5) на dt' и проинтегрируем в пределах от 0 до T :

$$\bar{I}(t') = \lim_{M \rightarrow \infty} \frac{1}{MT} \sum_{m=1}^M \int_0^T {}_m I(t') dt' = \lim_{M \rightarrow \infty} \frac{1}{MT} \int_0^{MT} I(t) dt, \quad (1.2-6)$$

а это то же самое, что и среднее во времени (1.2-4), если только последний предел существует.

1.3. ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ О НАЛОЖЕНИИ СЛУЧАЙНЫХ ВОЗМУЩЕНИЙ

Рассмотрим случай, когда точно K электронов поступают на анод в интервале длительностью T . Прежде чем интервал начнется, было предопределено прибытие этих K электронов в промежутке $(0, T)$, хотя каждый данный электрон может поступить на анод с одинаковой вероятностью в тот или другой момент времени. Подсчитываем все эти электроны от первого до K -го; следует подчеркнуть, что при этом подсчете не касаемся порядка поступления электронов на анод. Поэтому если t_k — момент поступления k -го электрона, то вероятность нахождения t_k в промежутке $(t, t+dt)$ равна $\frac{dt}{T}$.

T принято весьма большим сравнительно с интервалом значений t , для которого $F(t)$ значительно отличается от нуля. В физических задачах подобный интервал обычно существует; назовем его Δ , хотя он и не очень определен. Тогда если точно K электронов поступают на анод в промежутке $(0, T)$, то создаваемый ими эффект приблизительно равен

$$I_K(t) = \sum_{k=1}^K F(t-t_k), \quad (1.3-1)$$

причем степень приближения весьма хороша во всем интервале за исключением участка внутри Δ у конечных точек.

Допустим, что исследуется большое число M интервалов длительностью T . Число промежутков, в которые на анод поступает точно K электронов, будет равно в первом приближении $M\rho(K)$, где $\rho(K)$ находится из (1.1—3). Для фиксированного значения t и для каждого интервала, в который поступает K электронов, $I_K(t)$ будет иметь определенное значение. Когда $M \rightarrow \infty$, то среднее значение $I_K(t)$, найденное путем усреднения по интервалам, равно

$$\overline{I_K(t)} = \int_0^T \frac{dt_1}{T} \dots \int_0^T \frac{dt_K}{T} \sum_{k=1}^K F(t-t_k) = \sum_{k=1}^K \int_0^T \frac{dt_k}{T} F(t-t_k), \quad (1.3-2)$$

и если $\Delta < t < T - \Delta$, то

$$\overline{I_K(t)} = \frac{K}{T} \int_{-\infty}^{+\infty} F(t) dt. \quad (1.3-3)$$

Если теперь усредним $I(t)$ по всем M интервалам, а не только по тем, в которые на анод поступает K электронов, то получим,

если $M \rightarrow \infty$:

$$\begin{aligned} \overline{I(t)} &= \sum_{K=0}^{\infty} p(K) \overline{I_K(t)} = \sum_{K=0}^{\infty} \frac{K}{T} \frac{(\nu T)^K}{K!} e^{-\nu T} \int_{-\infty}^{+\infty} F(t) dt = \\ &= \nu \int_{-\infty}^{+\infty} F(t) dt, \end{aligned} \quad (1.3-4)$$

что и доказывает первую часть теоремы. Этот детальный вывод был применен для доказательства сравнительно простого соотношения (1.3-4), чтобы иллюстрировать метод, который может быть применен для доказательства более сложных выводов. Конечно, соотношение (1.3-4) легко установить, замечая, что интеграл представляет собой среднее значение эффекта, вызываемого поступлением одного электрона, причем это среднее значение взято за 1 сек., а ν есть среднее число поступлений электронов за 1 сек.

Чтобы доказать вторую часть (1.2-3) теоремы, сначала вычислим $\overline{I^2(t)}$ и воспользуемся соотношением

$$[\overline{I(t) - \overline{I(t)}}]^2 = \overline{I^2(t)} - 2\overline{I(t)\overline{I(t)}} + \overline{I(t)^2} = \overline{I^2(t)} - \overline{I(t)}^2. \quad (1.3-5)$$

Из определения $I_K(t)$ в уравнении (1.3-1) следует

$$I_K(t) = \sum_{k=1}^K \sum_{m=1}^K F(t-t_k) F(t-t_m).$$

Усредняя это выражение по всем значениям t_1, t_2, \dots, t_K , поддерживая t постоянным, как и в (1.3-2), получим

$$\overline{I_K^2(t)} = \sum_{k=1}^K \sum_{m=1}^K \int_0^T \frac{dt_1}{T} \dots \int_0^T \frac{dt_K}{T} F(t-t_k) F(t-t_m).$$

Кратный интеграл имеет два различных значения. Если $k=m$, то его значение равно

$$\int_0^T F^2(t-t_k) \frac{dt_k}{T},$$

а если $k \neq m$, то

$$\int_0^T F(t-t_k) \frac{dt_k}{T} \int_0^T F(t-t_m) \frac{dt_m}{T}$$

Подсчет числа членов в двойной сумме показывает, что первое значение имеют K членов, а второе значение ($K^2 - K$) членов.

Следовательно, если $\Delta < t < T - \Delta$, то

$$\overline{I_K^2(t)} = \frac{K}{T} \int_{-\infty}^{+\infty} F^2(t) dt + \frac{K(K-1)}{T^2} \left[\int_{-\infty}^{+\infty} F(t) dt \right]^2$$

Усредняя по всем интервалам, а не только по тем, в которые на анод поступает K электронов, получим

$$\overline{I^2(t)} = \sum_{K=0}^{\infty} p(K) \overline{I_K^2(t)} = \nu \int_{-\infty}^{+\infty} F^2(t) dt + \overline{I(t)^2},$$

где суммирование по K выполняется так же, как в (1.3—4), а после суммирования подставляется значение (1.3—4) для $\overline{I(t)}$. Сравнение с (1.3—5) доказывает вторую часть теоремы.

1.4. РАСПРЕДЕЛЕНИЕ ТОКА $I(t)$

Если выполняются некоторые условия, то часть времени, в течение которого ток $I(t)$ заключен в пределах $(I, I+dI)$, равна $P(I) dI$, где при $\nu \rightarrow \infty$ плотность вероятностей $P(I)$ сходится к

$$\frac{1}{\sigma_I \sqrt{2\pi}} e^{-\frac{(I-\bar{I})^2}{2\sigma_I^2}} \quad (1.4-1)$$

Здесь \bar{I} — среднее значение $I(t)$, равное (1.2—2), а квадрат стандартного отклонения σ_I , т. е. дисперсия, равна (1.2—3). Нормальное распределение следовало ожидать в силу «центральной предельной теоремы» теории вероятностей. Она утверждает, что при некоторых условиях распределение суммы большого числа случайных переменных сходится к нормальному закону распределения, дисперсия которого равна сумме дисперсий отдельных переменных. Подобным же образом среднее значение нормального распределения равно сумме средних значений отдельных переменных.

До сих пор говорилось о предельной форме плотности вероятностей $P(I)$. Можно написать точное выражение для $P(I)$, которое, однако, весьма сложно. Из точного выражения может быть получена предельная форма. Найдем теперь это выражение.

В соответствии с тем, как это было сделано при доказательстве теоремы о наложении случайных возмущений, ищем плотность вероятностей $P(I)$ для значений $I(t)$, наблюдаемых через t секунд от начала каждого из промежутков длительностью T , составляющих большое число M .

$$\begin{aligned} & \text{Вероятность нахождения } I(t) \text{ в интервале } (I, I+dI) = \\ & = \sum_{K=0}^{\infty} (\text{вероятность поступления точно } K \text{ электронов}) \times \\ & \times (\text{вероятность нахождения } I_K(t) \text{ в интервале } (I, I+dI) \\ & \text{при поступлении на анод точно } K \text{ электронов)}. \end{aligned}$$

Обозначая последнюю вероятность в суммировании через $P_K(I) dI$, используя применявшиеся ранее обозначения и отбрасывая множитель dI , получим

$$P(I) = \sum_{K=0}^{\infty} p(K) P_K(I). \quad (1.4-2)$$

Вычислим $P_K(I)$ методом «характеристических функций», используя определение $I_K(t)$

$$I_K(t) = \sum_{k=1}^K F(t-t_k). \quad (1.3-1)$$

Этот метод будет применен в его простейшей форме: вероятность того, что сумма K независимых случайных переменных

$$x_1 + x_2 + \dots + x_K$$

находится в интервале между X и $X+dX$, равна

$$dX \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iXu} \prod_{k=1}^K (\text{сред. знач. } e^{ix_k u}) du. \quad (1.4-3)$$

Среднее значение $e^{ix_k u}$, т. е. характеристическая функция распределения x_k , находится усреднением по значениям x_k . Хотя это и наиболее простая форма метода, но она также и наименее общая, так как интеграл в некоторых важных случаях расходится. Примером такого случая является распределение, которое дает вероятность $\frac{1}{2}$, что $x_k = -1$, и $\frac{1}{2}$, что $x_k = 1$. Однако в таких случаях все же формально можно пользоваться уравнением (1.4-3), применяя соотношение

$$\int_{-\infty}^{+\infty} e^{-iau} du = 2\pi\delta(a), \quad (1.4-4)$$

где $\delta(a) = 0$ за исключением $a=0$, когда $\delta(0) = \infty$, а ее интеграл, взятый в пределах от $a=-\epsilon$ до $a=+\epsilon$, равен единице (здесь $\epsilon > 0$):

Если заменить x_k на $F(t-t_k)$, то видно, что среднее значение $e^{ix_k u}$ равно

$$\frac{1}{T} \int_0^T \exp[iuF(t-t_k)] dt_k.$$

Все K характеристических функций одинаковы, и, следовательно, из (1.4-3) $P_K(I) dI$ равно

$$dI \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iIu} \left\{ \frac{1}{T} \int_0^T \exp[iuF(t-\tau)] d\tau \right\}^K du,$$

Хотя при выводе этого соотношения было принято $K > 0$, но оно также справедливо и для $K = 0$ [если только использовано (1.4—4)]. В этом случае $P_0(I) = \delta(I)$, поскольку $I = 0$, когда электроны не поступают на анод.

Вводя выражение для $P_K(I)$ и выражение (1.1—3) для $p(K)$ в уравнение (1.4—2) и выполняя суммирование, получим

$$P(I) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp \left\{ -iIu - \nu T + \nu \int_0^T \exp [iuF(t - \tau)] d\tau \right\} du. \quad (1.4-5)$$

Первая показательная функция может быть несколько упрощена. Воспользовавшись соотношением

$$\nu T = \nu \int_0^T d\tau$$

можем написать

$$-\nu T + \nu \int_0^T \exp [iuF(t - \tau)] d\tau = \nu \int_0^T \left\{ \exp [iuF(t - \tau)] - 1 \right\} d\tau.$$

Полагаем, что $\Delta < t < T - \Delta$, где Δ есть интервал, рассмотренный в связи с уравнением (1.3—1). Принимая $|F(t - \tau)| = 0$ при $|t - \tau| > \Delta$, можно написать последнее выражение в виде

$$\nu \int_{-\infty}^{+\infty} (e^{iuF(t)} - 1) dt. \quad (1.4-6)$$

Подставив это выражение в (1.4—5), получим искомое уравнение для $P(I)$

$$P(I) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp \left[-iIu + \nu \int_{-\infty}^{+\infty} (e^{iuF(t)} - 1) dt \right] du. \quad (1.4-7)$$

Представление о тех условиях, при которых плотность вероятностей сходится к нормальному закону (1.4—1), можно получить из (1.4—7), разлагая (1.4—6) в ряд по u и определяя, когда можно пренебречь членами с u^3 и с высшими степенями u . Это сделано в разделе 1.6 для другого, несколько более общего случая.

1.5. ОБОБЩЕНИЕ ТЕОРЕМЫ О НАЛОЖЕНИИ СЛУЧАЙНЫХ ВОЗМУЩЕНИЙ

Вместо выражения (1.2—1) для тока дробового эффекта $I(t)$ будем теперь иметь дело с током

$$I(t) = \sum_{k=-\infty}^{+\infty} a_k F(t - t_k), \quad (1.5-1)$$

где $F(t)$ — функция того же типа, что и ранее, и где... $a_1, a_2, \dots, a_k, \dots$ — независимые случайные переменные, все имеющие одинаковое распределение. Предполагается, что моменты $\overline{a^n}$ существуют и что все события являются случайными.

Обобщенная теорема утверждает, что n -й семиинвариант плотности вероятностей $P(I)$ для тока I (1.5—1) равен

$$\lambda_n = \nu \overline{a^n} \int_{-\infty}^{+\infty} [F(t)]^n dt, \quad (1.5-2)$$

где ν есть ожидаемое число событий в 1 сек.

Семиинварианты распределения определяются как коэффициенты в разложении

$$\log_e (\text{сред. } e^{iu}) = \sum_{n=1}^N \frac{\lambda_n}{n!} (iu)^n + \eta(u^N), \quad (1.5-3)$$

т. е. как коэффициенты в разложении логарифма характеристической функции. Семиинварианты λ связаны с моментами распределения. Поэтому если m_1, m_2, \dots обозначают моменты первого, второго и т. д. порядков относительно нуля, то

$$\text{сред. } e^{iu} = 1 + \sum_{n=1}^N \frac{m_n}{n!} (iu)^n + \eta(u^N).$$

Комбинируя это соотношение с выражением для λ , можно показать, что

$$\begin{aligned} \bar{I} &= m_1 = \lambda_1, \\ \bar{I}^2 &= m_2 = \lambda_2 + \lambda_1 m_1, \\ \bar{I}^3 &= m_3 = \lambda_3 + 2\lambda_2 m_1 + \lambda_1 m_2. \end{aligned}$$

Отсюда следует, что $\lambda_1 = \bar{I}$, а $\lambda_2 = (\bar{I} - \bar{I}^2)$. Следовательно, из (1.5—2) можно получить первоначальную формулировку теоремы, если положить n равным единице и двум и принять, что все a равны единице.

Обобщение теоремы немедленно следует из обобщения выражения (1.4—7) для плотности вероятностей. Прodelывая такие же операции, что и в разделе 1.4, и заменяя x_k на $a_k F(t - t_k)$, получим

$$\text{сред. } e^{ix_k u} = \frac{1}{T} \int_{-\infty}^{+\infty} q(a) da \int_0^T \exp [iua F(t - t_k)] dt_k,$$

где $q(a)$ есть функция плотности вероятностей для a . Следовательно,

плотность вероятностей $P(I)$ для тока I , заданного в виде (1.5—1), равна

$$P(I) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp \left[-iIu + \nu \int_{-\infty}^{+\infty} q(a) da \int_{-\infty}^{+\infty} (e^{iuaF(t)} - 1) dt \right] du. \quad (1.5-4)$$

Из (1.5—4) логарифм характеристической функции для $P(I)$ равен

$$\nu \int_{-\infty}^{+\infty} q(a) da \int_{-\infty}^{+\infty} (e^{iuaF(t)} - 1) dt = \sum_{n=1}^{\infty} \frac{(iu)^n}{n!} \nu \int_{-\infty}^{+\infty} q(a) da a^n \int_{-\infty}^{+\infty} F^n(t) dt.$$

Сравнение с рядом (1.5—3), определяющим семинварианты, приводит к обобщению теоремы, сформулированному в виде (1.5—2).

Могут быть сделаны и другие обобщения теоремы. Например, положим в выражении (1.5—1) для тока $I(t)$, что $t_1, t_2, \dots, t_k, \dots$, хотя и продолжают быть случайными переменными, но теперь не обязательно распределены в соответствии с принятыми ранее законами. Допустим, что дана плотность вероятностей $p(x)$, где x — промежуток между следующими друг за другом событиями:

$$t_2 = t_1 + x_1, \quad (1.5-5)$$

$$t_3 = t_2 + x_2 = t_1 + x_1 + x_2 \text{ и т. д.}$$

Для рассмотренного выше случая

$$p(x) = \nu e^{-\nu x}. \quad (1.5-6)$$

Предполагаем, что ожидаемое число событий в 1 сек. равно ν . Возьмем частный, но важный случай, для которого

$$\begin{aligned} F(t) &= 0, & t < 0, \\ F(t) &= e^{-at}, & t > 0. \end{aligned} \quad (1.5-7)$$

Для очень длинного промежутка, простирающегося от $t=t_1$ до $t=T+t_1$, внутри которого происходит точно K событий, будем иметь, если t не близко к концам промежутка:

$$\begin{aligned} I(t) &= a_1 F(t-t_1) + a_2 F(t-t_1-x_1) + \dots + a_{K+1} F(t-t_1-x_1-\dots-x_K) = \\ &= a_1 F(t') + a_2 F(t'-x_1) + \dots + a_{K+1} F(t'-x_1-\dots-x_K), \end{aligned}$$

$$\begin{aligned} I^2(t) &= a_1^2 F^2(t') + a_2^2 F^2(t'-x_1) + \dots + a_{K+1}^2 F^2(t'-x_1-\dots-x_K) + \\ &+ 2a_1 a_2 F(t') F(t'-x_1) + \dots + 2a_1 a_{K+1} F(t') F(t'-x_1-\dots-x_K) + \\ &+ 2a_2 a_3 F(t'-x_1) F(t'-x_1-x_2) + \dots + \dots \end{aligned}$$

где $t' = t - t_1$. Если проинтегрируем $I^2(t)$ по всему интервалу $0 < t' < T$ и опустим штрих, то получим приближенно

$$\int_0^T I^2(t) dt = (a_1^2 + \dots + a_{K+1}^2) \varphi(0) + \\ + 2a_1 a_2 \varphi(x_1) + 2a_1 a_3 \varphi(x_1 + x_2) + \dots + 2a_1 a_{K+1} \varphi(x_1 + \dots + x_K) + \\ + 2a_2 a_3 \varphi(x_2) + \dots + \dots + 2a_K a_{K+1} \varphi(x_K),$$

где

$$\varphi(x) = \int_{-\infty}^{+\infty} F(t) F(t-x) dx.$$

Если разделить обе части на T и положить, что K и T очень велики, то

$$\frac{K}{T} \frac{a_1^2 + \dots + a_{K+1}^2}{K} \varphi(0) \approx \sqrt{\bar{a}^2} \varphi(0) \\ \frac{1}{T} \left[a_1 a_2 \varphi(x_1) + a_2 a_3 \varphi(x_2) + \dots + a_K a_{K+1} \varphi(x_K) \right] = \\ = \frac{K}{T} \text{сред.} [a_k a_{k+1} \varphi(x_k)] \approx \\ \approx \sqrt{\bar{a}^2} \int_0^\infty \varphi(x) p(x) dx, \\ \frac{1}{T} \left[a_1 a_3 \varphi(x_1 + x_2) + \dots \right] = \frac{K-1}{T} \text{сред.} [a_k a_{k+2} \varphi(x_k + x_{k+1})] \approx \\ \approx \sqrt{\bar{a}^2} \int_0^\infty dx_1 \int_0^\infty dx_2 p(x_1) p(x_2) \varphi(x_1 + x_2).$$

Соответственно

$$\overline{I^2(t)} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T I^2(t) dt = \\ = \sqrt{\bar{a}^2} \varphi(0) + 2\sqrt{\bar{a}^2} \left[\int_0^\infty p(x) \varphi(x) dx + \right. \\ \left. + \int_0^\infty dx_1 \int_0^\infty dx_2 p(x_1) p(x_2) \varphi(x_1 + x_2) + \dots \right].$$

Для нашего частного случая экспоненциальной формы $F(t)$ имеем

$$\varphi(x) = \frac{e^{-ax}}{2a},$$

а кратные интегралы, встречающиеся в выражении для $\overline{I^2(t)}$, могут быть представлены при помощи

$$q = \int_0^{\infty} p(x) e^{-ax} dx. \quad (1.5-8)$$

Поэтому

$$2\alpha \overline{I^2(t)} = \nu \overline{a^2} + 2\overline{a^2} \nu \frac{q}{1-q},$$

а так как

$$\overline{I(t)} = \nu \overline{a} \int_{-\infty}^{+\infty} F(t) dt = \nu \frac{\overline{a}}{a},$$

то получим

$$\overline{I^2(t)} - \overline{I(t)}^2 = \frac{\nu \overline{a^2}}{2\alpha} + \left(\nu \frac{\overline{a}}{a}\right)^2 \left[\frac{\alpha q}{\nu(1-q)} - 1 \right]. \quad (1.5-9)$$

Уравнения (1.5-8) и (1.5-9) представляют собой обобщение теоремы при условии сохранения ограничений, рассмотренных в связи с уравнениями (1.5-5) и (1.5-7). Были сделаны и другие обобщения¹⁾, но здесь ограничимся этим.

Читатель может найти интересным проверку того, даст ли уравнение (1.5-9) правильный ответ, если $p(x)$ задано в виде (1.5-6), а также исследовать случай, когда события разделены равными промежутками.

1.6. СХОДИМОСТЬ РАСПРЕДЕЛЕНИЯ ТОКА I К НОРМАЛЬНОМУ ЗАКОНУ

В разделе 1.5 было показано, что плотность вероятностей $P(I)$ для тока шумов I может быть формально представлена в виде

$$P(I) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp \left[-iIu + \sum_{n=1}^{\infty} (iu)^n \frac{\lambda_n}{n!} \right] du, \quad (1.6-1)$$

где λ_n есть n -й семиинвариант, определяемый уравнением (1.5-2). Полагая

$$\begin{aligned} \lambda_2 &= \sigma^2, \\ x &= \frac{I - \lambda_1}{\sigma} = \frac{I - \overline{I}}{\sigma} \end{aligned} \quad (1.6-2)$$

и разлагая

$$\exp \sum_{n=3}^{\infty} (iu)^n \frac{\lambda_n}{n!}$$

¹⁾ См. Е. Н. Роулан, *Proc. Camb. Phil. Soc.*, 32, 580—597, 1936. Он распространяет теорему на случай существования двух функций вместо одной, обозначенной нами $I(t)$. А. Я. Хинчин в Известиях АН СССР (сер. мат., № 3, 1938) продолжил и уточнил работу Роуланда.

в ряд по степеням u , почленно интегрируя этот ряд, пользуясь соотношениями

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} (iu\sigma)^n \exp\left(-iu\sigma x - \frac{u^2\sigma^2}{2}\right) du = (-)^n \sigma^{-1} \varphi^{(n)}(x),$$

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

и, наконец, собирая члены согласно порядку степени $\nu^{-\frac{1}{2}}$, получим

$$P(I) \approx \sigma^{-1} \varphi^{(0)}(x) - \frac{\lambda_3 \sigma^{-4}}{3!} \varphi^{(3)}(x) +$$

$$+ \left[\frac{\lambda_4 \sigma^{-5}}{4!} \varphi^{(4)}(x) + \frac{\lambda_5 \sigma^{-7}}{72} \varphi^{(6)}(x) \right] + \dots \quad (1.6-3)$$

Первый член дает нормальное распределение, а прочие члены показывают, как достигается такое распределение, когда $\nu \rightarrow \infty$.

1.7. СОСТАВЛЯЮЩИЕ ФУРЬЕ ТОКА $I(t)$

В некоторых аналитических работах ток шумов представляется в виде

$$I(t) = \frac{a_0}{2} + \sum_{n=1}^N \left(a_n \cos \frac{2\pi n t}{T} + b_n \sin \frac{2\pi n t}{T} \right) \quad (1.7-1)$$

и в определенном месте работы полагается, что T и N стремятся к бесконечности. Коэффициенты a_n и b_n , $1 \leq n \leq N$, рассматриваются как независимые случайные переменные, распределенные вокруг нуля по нормальному закону.

В соответствии с нашим обычным подходом к дробовому эффекту предположим, что в течение промежутка $(0, T)$ на анод поступают точно K электронов, так что ток шумов в этом промежутке равен

$$I_K(t) = \sum_{k=1}^K F(t-t_k). \quad (1.7-2)$$

Коэффициентами разложения $I_K(t)$ в ряд Фурье в интервале $(0, T)$ являются a_{nK} и b_{nK} , где

$$a_{nK} - ib_{nK} = \frac{2}{T} \sum_{k=1}^K \int_0^T F(t-t_k) \exp\left(-i \frac{2\pi n t}{T}\right) dt \approx$$

$$\approx \frac{2}{T} \sum_{k=1}^K \int_{-\infty}^{+\infty} F(t) \exp\left[-i \frac{2\pi n}{T} (t+t_k)\right] dt = R_n e^{-i\varphi_n} \sum_{k=1}^K e^{-in\theta_k}. \quad (1.7-3)$$

В этом выражении

$$\theta_k = \frac{2\pi t_k}{T}, \quad (1.7-4)$$

$$R_n e^{-i\varphi_n} = C_n - iS_n = \frac{2}{T} \int_{-\infty}^{+\infty} F(t) e^{-i2\pi nt/T} dt.$$

В предыдущих разделах моменты поступления электронов t_1, t_2, \dots, t_K рассматривались как K независимых случайных переменных, каждое из которых распределено равномерно в интервале $(0, T)$. Следовательно, и θ_k могут считаться случайными переменными, равномерно распределенными по интервалу от 0 до 2π .

Попутно заметим, что в уравнение (1.7-3) входит сумма из K случайно ориентированных единичных векторов. Когда K становится очень большим, как это будет в случае $\nu \rightarrow \infty$, то известно, что действительная и мнимая части этой суммы представляют собой случайные переменные, которые стремятся стать независимыми и нормально распределенными вокруг нуля. Это дает представление о том, каким образом появляется нормальное распределение коэффициентов. Усреднение по θ_k в уравнении (1.7-3) дает, когда $n > 0$:

$$\bar{a}_{nK} = \bar{b}_{nK} = 0. \quad (1.7-5)$$

После дальнейших алгебраических выкладок

$$\begin{aligned} \bar{a}_{nK}^2 &= \bar{b}_{nK}^2 = \frac{K}{2} R_n^2, \\ \overline{a_{nK} b_{nK}} &= \overline{a_{nK} a_{mK}} = \overline{b_{nK} b_{mK}} = 0, \end{aligned} \quad (1.7-6)$$

где $n \neq m$ и $n, m > 0$.

До сих пор рассматривался случай поступления на анод точно K электронов в промежутке длительностью T . Теперь переходим к общему случаю поступления любого числа электронов, используя формулы, аналогичные

$$\bar{a}_n^2 = \sum_{K=0}^{\infty} p(K) \bar{a}_{nK}^2, \quad (1.7-7)$$

как это было сделано в разделе 1.3. Поэтому при $n > 0$

$$\begin{aligned} \bar{a}_n &= \bar{b}_n = 0, \\ \bar{a}_n^2 &= \bar{b}_n^2 = \frac{\sqrt{T}}{2} R_n^2 = \sigma_n^2, \\ \overline{a_n b_n} &= \overline{a_n a_m} = \overline{b_n b_m} = 0, \quad n \neq m. \end{aligned} \quad (1.7-8)$$

Во второй строчке σ_n обозначает стандартное отклонение величин a_n и b_n . Можно получить выражение для σ_n^2 в несколько иной форме,

написав

$$f_n = \frac{n}{T} = n\Delta f, \quad \Delta f = \frac{1}{T}, \quad (1.7-9)$$

где f_n — частота n -й составляющей. Воспользовавшись (1.7-4), получим

$$\sigma_n^2 = 2\nu\Delta f \left| \int_{-\infty}^{+\infty} F(t)e^{-i2\pi f_n t} dt \right|^2 \quad (1.7-10)$$

Поэтому σ_n^2 пропорционально $\frac{\nu}{T}$.

Функция плотности вероятностей $P(a_1, \dots, a_N, b_1, \dots, b_N)$ для $2N$ коэффициентов $a_1, \dots, a_N, b_1, \dots, b_N$ может быть выведена подобно плотности вероятностей тока шумов в разделе 1.4. Здесь N произвольно, но фиксировано. Выражение, аналогичное (1.4-5), есть интеграл кратности $2N$

$$P(a_1, \dots, a_N, b_1, \dots, b_N) = \\ = (2\pi)^{-2N} \int_{-\infty}^{+\infty} du_1 \dots \int_{-\infty}^{+\infty} dv_N \exp[-i(a_1 u_1 + \dots + b_N v_N) - \nu T + \nu T E], \quad (1.7-11)$$

где

$$E = \frac{1}{2\pi} \int_0^{2\pi} d\theta \exp \left[i \sum_{n=1}^N (u_n C_n + v_n S_n) \cos n\theta + (v_n C_n - u_n S_n) \sin n\theta \right]. \quad (1.7-12)$$

В этом уравнении $C_n - iS_n$ представляет собой преобразование Фурье (1.7-4) функции $F(t)$.

Вслед за этим нужно показать, что (1.7-11) сходится к нормальному закону в $2N$ измерениях, когда $\nu \rightarrow \infty$. Эта задача оказывается весьма сложной. Представление о том, как выглядит это выражение, можно получить, рассматривая частный случай, когда $F(t)$ является четной функцией t , и пренебрегая некоторыми членами. Тогда

$$P(a_1, \dots, a_N, b_1, \dots, b_N) = (1 + \eta) \prod_{n=1}^N \frac{e^{-(x_n^2 + y_n^2)/2}}{2\pi\sigma_n^2}, \quad (1.7-13)$$

где

$$x_n = \frac{a_n}{\sigma_n}, \quad y_n = \frac{b_n}{\sigma_n}, \quad (1.7-14)$$

$$\eta = (2\nu T)^{-1/2} \sum_{k,l} [x_{k+l}(x_k x_l - y_k y_l) + 2y_{k+l} y_k y_l],$$

а суммирование производится в пределах $2 \leq (k+l) \leq N$ при $k \leq l$.

Отсюда видно, что если T и N поддерживаются постоянными, то корректирующий член η стремится к нулю, когда ν становится очень большим. Весьма грубое представление о порядке величины η получают, считая, что x и y можно заменить единицами. Полагая далее, что сумма состоит из N^2 членов, каждый из которых может быть как положительным, так и отрицательным, получим, что величина суммы имеет порядок N . Тогда можно ожидать, что η имеет порядок $N(2\nu T)^{-1}$.

Г л а в а II

ЭНЕРГЕТИЧЕСКИЙ СПЕКТР И ФУНКЦИЯ КОРРЕЛЯЦИИ

В разделе 2.1 энергетический спектр и функция корреляции рассматриваемой функции времени, например, заданной в виде кривой, простирающейся до $t=\infty$, определяются соответственно уравнениями (2.1—3) и (2.1—4). Связь этих величин с формулами преобразования Фурье (2.1—5) и (2.1—6) вначале утверждается без доказательства; рассмотрение способа доказательства отнесено к разделам 2.3 и 2.4.

В разделе 2.3 рассмотрение основано на рядах Фурье, а в разделе 2.4 аналогичные результаты получаются более прямым путем на основе интегральной теоремы Парсеваля.

Если анализируемая функция содержит постоянную или периодические составляющие, то выводы раздела 2.1 должны быть дополнены, что и сделано в разделе 2.2.

Первые четыре раздела посвящены анализу заданной функции времени. Однако большинство приложений метода относится к функциям, которые ведут себя как более или менее случайные функции.

В математическом анализе подобная случайность обуславливается предположением, что функция t является также и функцией некоторых параметров, которые затем считаются случайными переменными. Этот вопрос разобран в разделе 2.5.

В разделе 2.6 выводы раздела 2.5 применяются для определения среднего энергетического спектра и средней функции корреляции тока дробового эффекта.

То же самое сделано в 2.7 для прямоугольной волны, полупериоды которой имеют случайную длительность. Пример, в котором интервалы предполагаются одинаковой длительности, но знак волны случаен, также рассмотрен в 2.7.

Представление тока шумов в виде тригонометрического ряда с коэффициентами, рассматриваемыми как случайные переменные, разбирается в разделе 2.8.

Последние два раздела 2.9 и 2.10 посвящены некоторым вопросам теории вероятностей, в них соответственно рассмотрен нормальный закон и центральная предельная теорема.

2.1. НЕКОТОРЫЕ РЕЗУЛЬТАТЫ ОБОБЩЕННОГО ГАРМОНИЧЕСКОГО АНАЛИЗА

Вначале сформулируем выводы, которые нужно получить, а затем покажем их правдоподобие при помощи методов, которые являются скорее эвристическими, чем строгими.

Допустим, что $I(t)$ есть одна из функций, которые упоминались выше. Можно представлять себе, что она задана в виде кривой, простирающейся от $t = -\infty$ до $t = \infty$. Можно также считать, что функция $I(t)$ составлена из большого числа синусоидальных составляющих, частоты которых лежат в диапазоне от 0 до $+\infty$. Это не обязательно должен быть ток шумов, но если считать эту функцию током шумов, то, протекая по сопротивлению 1 ом, этот ток рассеет некоторую среднюю мощность, скажем ρ вт. Часть этой средней мощности, выделяемой составляющими, частоты которых заключены между f и $f+df$, будем в дальнейшем обозначать $\omega(f) df$, следовательно:

$$\rho = \int_0^{\infty} \omega(f) df. \quad (2.1-1)$$

$\omega(f)$ имеет размерность энергии и на этом основании часто называется «энерго-частотным спектром» тока $I(t)$. В дальнейшем будем называть $\omega(f)$ просто «энергетическим спектром»¹⁾.

Математическая формулировка этих рассуждений приводит к совершенно четкому определению $\omega(f)$.

Пусть $\Phi(t)$ будет функция времени t , равная нулю вне интервала $0 \leq t \leq T$ и равная $I(t)$ внутри этого интервала. Ее спектр представляется следующим выражением:

$$S(f) = \int_0^T I(t) e^{-2\pi i f t} dt. \quad (2.1-2)$$

Энергетический спектр $\omega(f)$ определяется как

$$\omega(f) = \lim_{T \rightarrow \infty} \frac{2|S(f)|^2}{T} \quad (2.1-3)$$

где учитываются только значения $f > 0$ и предполагается, что этот предел существует. Это определение $\omega(f)$ применимо, когда $I(t)$ не имеет периодических членов и постоянной составляющей. В противном случае уравнение (2.1-3) должно быть либо дополнено, либо применен другой метод исследования. Эти вопросы будут разобраны в разделе 2.2.

¹⁾ В литературе для $\omega(f)$ принято также название «спектральная плотность мощности» или просто «спектральная плотность». (Прим. ред.)

Функция корреляции $\psi(\tau)$ для $I(t)$ определяется пределом

$$\psi(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T I(t) I(t + \tau) dt, \quad (2.1-4)$$

который предполагается существующим. $\psi(\tau)$ тесно связана с коэффициентами корреляции, применяемыми в статистической теории для измерения корреляции двух случайных переменных. В данном случае значение $I(t)$ в момент времени t есть одна переменная, а ее значение в другой момент времени $t + \tau$ есть другая переменная.

Энергетический спектр $\omega(f)$ и функция корреляции $\psi(\tau)$ связаны между собой следующими соотношениями¹⁾:

$$\omega(f) = 4 \int_0^{\infty} \psi(\tau) \cos 2\pi f \tau d\tau, \quad (2.1-5)$$

$$\psi(\tau) = \int_0^{\infty} \omega(f) \cos 2\pi f \tau df. \quad (2.1-6)$$

Отсюда видно, что $\psi(\tau)$ является четной функцией τ и что

$$\psi(0) = \rho. \quad (2.1-7)$$

Когда известна либо $\omega(f)$, либо $\psi(\tau)$, вторая функция может быть найдена, если только соответствующий интеграл является сходящимся.

2.2. ЭНЕРГЕТИЧЕСКИЙ СПЕКТР ДЛЯ ПОСТОЯННОЙ И ПЕРИОДИЧЕСКОЙ СОСТАВЛЯЮЩИХ

Как указывалось в разделе 2.1, если $I(t)$ имеет постоянную или периодическую составляющую, то предел в уравнении (2.1-3) для $f=0$ или для частоты периодической составляющей не существует. Пожалуй, наиболее удовлетворительным способом преодоления этой трудности с математической точки зрения является переход к оперированию с интегралом энергетического спектра

$$\int_0^f \omega(g) dg, \quad (2.2-1)$$

вместо того чтобы иметь дело с самим спектром $\omega(f)$.

Определение $\psi(\tau)$ в виде (2.1-4) сохраняется. Если, например:

$$I(t) = A + C \cos(2\pi f_0 t - \varphi), \quad (2.2-2)$$

¹⁾ Эти соотношения получены А. Я. Хинчиным в работе «Теория корреляции стационарных стохастических процессов» в 1934 г. (Прим. ред.)

то по уравнению (2.1—4) $\psi(\tau)$ равна

$$\psi(\tau) = A^2 + \frac{C^2}{2} \cos 2\pi f_0 \tau. \quad (2.2-3)$$

Формулы перехода (2.1—5) и (2.1—6) дают

$$\int_0^f \omega(g) dg = \frac{2}{\pi} \int_0^\infty \psi(\tau) \frac{\sin 2\pi f \tau}{\tau} d\tau, \quad (2.2-4)$$

$$\psi(\tau) = \int_0^\infty \cos 2\pi f \tau d \left[\int_0^f \omega(g) dg \right],$$

где последний интеграл должен рассматриваться как интеграл Стильтьеса. Если выражение (2.2—3) для $\psi(\tau)$ подставить в первую формулу (2.2—4), то получим

$$\int_0^f \omega(g) dg = \begin{cases} A^2, & \text{если } 0 < f < f_0, \\ A^2 + \frac{C^2}{2}, & \text{если } f > f_0. \end{cases} \quad (2.2-5)$$

Если это выражение использовать во второй формуле (2.2—4), то приращения дифференциала, очевидно, будут A^2 при $f=0$ и $\frac{C^2}{2}$ при $f=f_0$. Полученное выражение для $\psi(\tau)$ совпадает с первоначальным.

Теперь воспользуемся менее строгим, но более удобным методом рассмотрения случая периодических составляющих. Исследуя интеграл в выражении (2.2—5) для $\omega(f)$, можем написать

$$\omega(f) = 2A^2 \delta(f) + \frac{C^2}{2} \delta(f - f_0), \quad (2.2-6)$$

где $\delta(x)$ есть четная единичная импульсная функция, так что если $\epsilon > 0$, то

$$\int_0^\epsilon \delta(x) dx = \frac{1}{2} \int_{-\epsilon}^\epsilon \delta(x) dx = \frac{1}{2}, \quad (2.2-7)$$

а $\delta(x)=0$, за исключением $x=0$, когда $\delta(0)=\infty$. Это позволяет воспользоваться более простыми формулами перехода раздела 2.1. Сразу видно, что вторая из них (2.1—6) дает правильное выражение для $\psi(\tau)$. Первая формула (2.1—5) дает правильное выражение для $\omega(f)$, если интерпретировать интегралы следующим образом:

$$\int_0^\infty \cos 2\pi f \tau d\tau = \frac{1}{2} \delta(f), \quad (2.2-8)$$

$$\int_0^\infty \cos 2\pi f_0 \tau \cos 2\pi f \tau d\tau = \frac{1}{4} \delta(f - f_0).$$

Нетрудно показать, что это находится в согласии с основным представлением

$$\int_{-\infty}^{+\infty} e^{-i2\pi ft} dt = \int_{-\infty}^{+\infty} e^{i2\pi ft} dt = \delta(f), \quad (2.2-9)$$

которое, в свою очередь, следует из формального применения формулы интеграла Фурье и соотношения

$$\int_{-\infty}^{+\infty} \delta(f) e^{i2\pi ft} df = \int_{-\infty}^{+\infty} \delta(f) e^{-i2\pi ft} df = 1. \quad (2.2-10)$$

Следует помнить, что в (2.2-8) $f_0 > 0$, а $f \geq 0$, так что

$$\delta(f + f_0) = 0 \text{ для } f \geq 0.$$

Определение $\omega(f)$ в виде (2.1-3) дает сплошной участок энергетического спектра. Чтобы получить часть спектра, соответствующую постоянной и периодическим составляющим, примером которой служит уравнение (2.2-6) для $\omega(f)$ с функциями δ , надо дополнить (2.1-3) членами типа

$$\begin{aligned} & 2A^2 \delta(f) + \frac{C^2}{2} \delta(f - f_0) = \\ & = \left[\lim_{T \rightarrow \infty} \frac{2|S(0)|^2}{T^2} \right] \delta(f) + \left[\lim_{T \rightarrow \infty} \frac{2|S(f_0)|^2}{T^2} \right] \delta(f - f_0). \end{aligned} \quad (2.2-11)$$

Правильность этого выражения может быть проверена путем вычисления $S(f)$ для тока $I(t)$, заданного в виде (2.2-2), и нахождения пределов.

2.3. ОБСУЖДЕНИЕ ВЫВОДОВ ПЕРВОГО РАЗДЕЛА — РЯДЫ ФУРЬЕ

Тот факт, что соотношение между энергетическим спектром $\omega(f)$ и функцией корреляции $\varphi(\tau)$ задается формулами преобразования Фурье, непосредственно связан с теоремами Парсеваля для рядов и интеграла Фурье. Начнем с рядов Фурье и используем представление о рассеивании доли энергии каждой составляющей независимо от поведения других составляющих.

Пусть часть тока $I(t)$, приходящаяся на интервал $0 \leq t \leq T$, разложена в ряд Фурье

$$I(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos \frac{2\pi n t}{T} + b_n \sin \frac{2\pi n t}{T} \right) \quad (2.3-1)$$

где

$$a_n = \frac{2}{T} \int_0^T I(t) \cos \frac{2\pi n t}{T} dt, \quad (2.3-2)$$

$$b_n = \frac{2}{T} \int_0^T I(t) \sin \frac{2\pi n t}{T} dt.$$

Тогда для интервала $-\tau \leq t < T - \tau$

$$I(t + \tau) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos \frac{2\pi n(t+\tau)}{T} + b_n \sin \frac{2\pi n(t+\tau)}{T} \right) \quad (2.3-3)$$

Перемножая ряды для $I(t)$ и $I(t+\tau)$ и интегрируя по t , получим после некоторых сокращений

$$\begin{aligned} & \frac{1}{T} \int_0^T I(t) I(t + \tau) dt = \\ & = \frac{a_0^2}{4} + \sum_{n=1}^{\infty} \frac{1}{2} (a_n^2 + b_n^2) \cos \frac{2\pi n}{T} \tau + \zeta \left(\frac{\tau^2}{T} \right), \end{aligned} \quad (2.3-4)$$

где последний член является корректирующим и должен быть добавлен вследствие того, что ряд (2.3-3) не представляет $I(t+\tau)$ в промежутке $(T-\tau, T)$, если $\tau > 0$, или в промежутке $(0, -\tau)$, если $\tau < 0$.

Если ток $I(t)$ протекает по сопротивлению 1 ом в интервале $(0, T)$, то каждая составляющая рассеет некоторую среднюю мощность. Эта средняя мощность, выделенная составляющей с частотой $f_n = \frac{n}{T} \text{ гц}$, должна быть равна, как это следует из теории рядов Фурье и элементарных принципов:

$$\begin{aligned} & \frac{1}{2} (a_n^2 + b_n^2) \text{ вт}, \quad n \neq 0, \\ & \frac{a_0^2}{4} \text{ вт}, \quad n = 0. \end{aligned} \quad (2.3-5)$$

Ширина полосы, связанная с n -й составляющей, есть разность по частоте между $(n+1)$ -й и n -й составляющими

$$f_{n+1} - f_n = \frac{n+1}{T} - \frac{n}{T} = \frac{1}{T} \text{ гц}.$$

Следовательно, если среднюю мощность в полосе $(f, f+df)$ обозначить как $\omega(f) df$, то средняя мощность в полосе $f_{n+1} - f_n$ равна

$$\omega(f_n) (f_{n+1} - f_n) = \omega \left(\frac{n}{T} \right) \frac{1}{T},$$

а из (2.3—5) следует

$$\begin{aligned} \omega\left(\frac{n}{T}\right) \frac{1}{T} &= \frac{1}{2} (a_n^2 + b_n^2), & n \neq 0, \\ \omega(0) \frac{1}{T} &= \frac{a_0^2}{4}, & n = 0. \end{aligned} \quad (2.3-6)$$

Если коэффициенты в (2.3—4) заменить их значениями, выраженными через $\omega(f)$, то получим

$$\begin{aligned} \frac{1}{T} \int_0^T I(t)I(t+\tau) dt + \zeta\left(\frac{\tau I^2}{T}\right) &= \frac{1}{T} \sum_{n=0}^{\infty} \omega\left(\frac{n}{T}\right) \cos \frac{2\pi n\tau}{T} = \\ &= \int_0^{\infty} \omega\left(\frac{n}{T}\right) \cos \frac{2\pi n\tau}{T} \frac{dn}{T} = \int_0^{\infty} \omega(f) \cos 2\pi f\tau df, \end{aligned} \quad (2.3-7)$$

где полагаем T настолько большим, а $\omega(f)$ такого характера, что суммирование может быть заменено интегрированием.

Если I остается конечным, а $T \rightarrow \infty$ при τ , поддерживаемым постоянным, то корректирующий член слева становится ничтожно малым. Пользуясь определением (2.1—4) для функции корреляции $\psi(\tau)$, получим вторую из основных формул преобразования (2.1—6). Первая формула может быть сразу получена отсюда применением к $\omega(f)$ формулы двойного интеграла Фурье.

Кстати, соотношение (2.3—6) между $\omega(f)$ и коэффициентами a_n и b_n находится в согласии с определением $\omega(f)$ по формуле (2.1—3), как предела, содержащего $|S(f)|^2$. Из формулы (2.3—2) для a_n и b_n спектр $S(f_n)$ по уравнению (2.1—2) равен

$$S(f_n) = \frac{T}{2} (a_n - ib_n).$$

Тогда согласно (2.1—3) $\omega(f_n)$ равен следующему пределу при $T \rightarrow \infty$:

$$\frac{2}{T} |S(f_n)|^2 = \frac{2}{T} \frac{T^2}{4} (a_n^2 + b_n^2) = \frac{T}{2} (a_n^2 + b_n^2),$$

а это и есть выражение для $\omega\left(\frac{n}{T}\right)$ согласно (2.3—6).

2.4. ОБСУЖДЕНИЕ ВЫВОДОВ ПЕРВОГО РАЗДЕЛА — ТЕОРЕМА ПАРСЕВАЛЯ

Применение теоремы Парсеваля¹⁾ позволяет получить результаты раздела 2.1 более прямым путем, чем это дает метод, исполь-

¹⁾ Титчмарш, «Введение в теорию интегралов Фурье», Гостехиздат, 1948. (Прим. ред.)

зованный в предыдущем разделе. Эта теорема утверждает, что

$$\int_{-\infty}^{+\infty} F_1(f)F_2(f) df = \int_{-\infty}^{+\infty} G_1(t)G_2(-t) dt, \quad (2.4-1)$$

где F_1 , G_1 и F_2 , G_2 суть пары преобразований Фурье

$$F(f) = \int_{-\infty}^{+\infty} G(t) e^{-i2\pi ft} dt, \\ G(t) = \int_{-\infty}^{+\infty} F(f) e^{i2\pi ft} df. \quad (2.4-2)$$

Эти соотношения могут быть доказаны формальным образом подстановкой F_1 в виде интеграла, содержащего $G_1(t)$, в левую часть уравнения (2.4—1). Изменение порядка интегрирования и использование второй формулы (2.4—2) для замены F_2 на G_2 дает правую часть уравнения.

Положим теперь $G_1(t)$ и $G_2(t)$ равными нулю, за исключением интервалов длительностью T . Эти интервалы и соответствующие значения G_1 и G_2 составляют

$$G_1(t) = I(t), \quad 0 < t < T, \\ G_2(t) = I(-t + \tau), \quad \tau - T < t < \tau. \quad (2.4-3)$$

Из (2.4—3) следует, что $F_1(t)$ есть спектр $S(f)$ для $I(t)$, определяемый уравнением (2.1—2). Так как $I(t)$ вещественно, то из первого уравнения (2.4—2) следует, что

$$S(-f) = S^*(f), \quad (2.4-4)$$

где звездочка обозначает сопряженный комплекс, и, следовательно, $|S(f)|^2$ есть четная функция f .

Из первого уравнения (2.4—2) следует также

$$F_2(f) = \int_{\tau-T}^{\tau} I(-t + \tau) e^{-i2\pi ft} dt = \\ = \int_0^T I(t) e^{i2\pi f(t-\tau)} dt = S^*(f) e^{-i2\pi f\tau} \quad (2.4-5)$$

Если эти значения G и F подставить в (2.4—1), то получим

$$\int_{-\infty}^{+\infty} |S(f)|^2 e^{-i2\pi f\tau} df = \int_0^{T-\tau} I(t)I(t+\tau) dt, \quad (2.4-6)$$

Где использовано то обстоятельство, что $G_2(-t)$ повсюду равна нулю, за исключением интервала $-\tau < t < T - \tau$, и положено $\tau > 0$.

Если $\tau < 0$, то пределы интегрирования в (2.4—6) справа должны быть $-\tau$ и T .

Так как $|S(f)|^2$ есть четная функция f , то можно написать (2.4—6) в виде

$$\frac{1}{T} \int_0^T I(t)I(t + \tau) dt + \zeta \left(\frac{\tau I^2}{T} \right) = \int_0^{\infty} \frac{2|S(f)|^2}{T} \cos 2\pi f\tau df. \quad (2.4-7)$$

Если теперь определить функцию корреляции $\psi(\tau)$ как предел левой части уравнения при $T \rightarrow \infty$, а $\omega(f)$ как функцию

$$\omega(f) = \lim_{T \rightarrow \infty} \frac{2|S(f)|^2}{T}, \quad f > 0, \quad (2.1-3)$$

то будет получена вторая из основных формул преобразования (2.1—6). Как и раньше, первая может быть всегда получена при помощи интегральной теоремы Фурье.

Чтобы подойти к интерпретации $\omega(f)df$ как средней мощности, рассеиваемой в сопротивлении 1 ом теми составляющими $I(t)$, которые лежат в полосе частот $(f, f+df)$, надо положить в (2.4—7) $\tau=0$

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T I^2(t) dt = \int_0^{\infty} \omega(f) df. \quad (2.4-8)$$

Выражение слева представляет собой, очевидно, полную среднюю мощность, рассеиваемую в сопротивлении 1 ом, а правая часть — суммирование по всему диапазону частот от 0 до ∞ . Поэтому естественно истолковать $\omega(f)df$ как мощность, выделяемую составляющими в полосе частот $(f, f+df)$.

В предыдущих разделах речь шла об энергетических спектрах $\omega(f)$ и функции корреляции $\psi(\tau)$ для весьма общего типа функций. Следует заметить, что знание $\omega(f)$ не дает возможности определить первоначальную функцию $I(t)$. При нахождении $\omega(f)$, как это можно видеть из (2.1—3) или из (2.3—6), данные, связанные с фазовыми углами различных составляющих $I(t)$, исчезают. Действительно, как это видно из представления $I(t)$ в виде ряда Фурье (2.3—1) и из (2.3—6), можно найти бесконечное число различных функций, имеющих одинаковый спектр $\omega(f)$ и, следовательно, одинаковую функцию корреляции $\psi(\tau)$.

2.5. ГАРМОНИЧЕСКИЙ АНАЛИЗ СЛУЧАЙНЫХ ФУНКЦИЙ

Во многих приложениях теории, рассмотренной в предыдущих разделах, $I(t)$ есть функция t , обладающая определенной долей случайности. Например, $I(t)$ может быть кривая составляющей

скорости воздуха позади сетки, помещенной в аэродинамической трубе, или, что наиболее интересно, кривая тока шумов.

В некоторых математических работах эта случайность определяется введением в $I(t)$ ряда параметров, которые затем считаются случайными переменными. Так, в дробовом эффекте моменты поступления электронов t_1, t_2, \dots, t_k приняты за параметры и каждый предполагается равномерно распределенным в интервале $(0, T)$.

Для любого данного ряда значений параметров $I(t)$ имеет определенные энергетический спектр $\omega(f)$ и функцию корреляции $\psi(\tau)$. Однако теперь основной интерес представляют те функции, которые дают средние значения $\omega(f)$ и $\psi(\tau)$ для фиксированных f и τ . Эти функции можно найти путем усреднения $\omega(f)$ и $\psi(\tau)$ по интервалу изменения параметров, пользуясь, конечно, функциями распределения параметров.

Усредняя обе части соответствующих уравнений в разделах 2.1 и 2.2, видим, что основные формулы преобразования (2.1—5) и (2.1—6) остаются неизменными. Поэтому

$$\bar{\omega}(f) = 4 \int_0^{\infty} \bar{\psi}(\tau) \cos 2\pi f \tau d\tau, \quad (2.5-1)$$

$$\bar{\psi}(\tau) = \int_0^{\infty} \bar{\omega}(f) \cos 2\pi f \tau df, \quad (2.5-2)$$

где черточки указывают на усреднение по параметрам при f или τ , сохраняемым постоянным. Определения $\bar{\omega}$ и $\bar{\psi}$ в этих уравнениях можно также получить из (2.1—3) и (2.1—4)

$$\bar{\omega}(f) = \lim_{T \rightarrow \infty} \frac{2|S(f)|^2}{T} \quad (2.5-3)$$

$$\bar{\psi}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \overline{I(t)I(t+\tau)} dt. \quad (2.5-4)$$

При усреднении по параметрам значения t и τ сохраняются постоянными. В уравнении (2.5—3) $S(f)$, рассматриваемая как функция параметров, связана с $I(t)$ следующим соотношением:

$$S(f) = \int_0^T I(t) e^{-2\pi i f t} dt. \quad (2.1-2)$$

Подобные же выражения могут быть получены для среднего энергетического спектра постоянной и периодической составляющих. Все, что нужно сделать, — это усреднить выражение (2.2—11).

Иногда среднее значение произведения $I(t)I(t+\tau)$ в форме (2.5—4) для $\overline{\psi}(\tau)$ не зависит от времени T . Это позволяет сразу выполнить интегрирование и получить

$$\overline{\psi}(\tau) = \overline{I(t)I(t+\tau)}. \quad (2.5—5)$$

При этом достигается значительное упрощение и, кажется, простейшим способом вычисления $\overline{\omega}(f)$ для $I(t)$ такого типа является нахождение сначала $\overline{\psi}(\tau)$, а затем применение формулы перехода (2.5—1).

2.6. ПЕРВЫЙ ПРИМЕР — ДРОБОВОЙ ЭФФЕКТ

Вычислим сначала среднее значение правой части уравнения (2.5—5). Используя метод усреднения, широко применявшийся в главе I, найдем

$$\overline{I(t)I(t+\tau)} = \sum_{K=0}^{\infty} \rho(K) \overline{I_K(t)I_K(t+\tau)}, \quad (2.6—1)$$

где $\rho(K)$ есть вероятность поступления на анод точно K электронов в промежутке $(0, T)$

$$\rho(K) = \frac{(\nu T)^K}{K!} e^{-\nu T} \quad (1.1—3)$$

а

$$I_K(t) = \sum_{k=1}^K F(t - t_k). \quad (1.3—1)$$

Перемножая $I_K(t)$ и $I_K(t+\tau)$ и усредняя t_1, t_2, \dots, t_K по их интервалам, получим

$$\overline{I_K(t)I_K(t+\tau)} = \sum_{k=1}^K \sum_{m=1}^K \int_0^T \frac{dt_1}{T} \dots \int_0^T \frac{dt_K}{T} F(t - t_k) F(t + \tau - t_m).$$

Это уравнение совершенно подобно выражению для $\overline{I_K^2(t)}$, которое применялось в разделе 1.3 для доказательства теоремы о наложении случайных возмущений, и может трактоваться таким же точно образом. Поэтому, если t и $t+\tau$ лежат между Δ и $T-\Delta$, то написанное выше выражение приобретает вид

$$\frac{K}{T} \int_{-\infty}^{+\infty} F(t)F(t+\tau) dt + \frac{K(K-1)}{T^2} \left[\int_{-\infty}^{+\infty} F(t) dt \right]^2$$

Если его подставить в (2.6—1) и выполнить суммирование, то получим выражение, не зависящее от T . Поэтому можно воспользоваться (2.5—5) и найти

$$\overline{\psi}(\tau) = \nu \int_{-\infty}^{+\infty} F(t)F(t+\tau) dt + \overline{I(t)}^2, \quad (2.6—2)$$

где было применено следующее уравнение для среднего значения тока:

$$\overline{I(t)} = \nu \int_{-\infty}^{+\infty} F(t) dt. \quad (1.3-4)$$

Чтобы вычислить $\overline{\omega(f)}$ из $\overline{\psi(\tau)}$, удобно воспользоваться тем обстоятельством, что $\psi(\tau)$ есть всегда четная функция τ , и, следовательно, (2.5—1) можно также написать в виде

$$\overline{\omega(f)} = 2 \int_{-\infty}^{+\infty} \overline{\psi(\tau)} \cos 2\pi f\tau d\tau. \quad (2.6-3)$$

Тогда

$$\begin{aligned} \overline{\omega(f)} &= 2\nu \int_{-\infty}^{+\infty} dt F(t) \int_{-\infty}^{+\infty} d\tau F(t+\tau) \cos 2\pi f\tau + 2 \int_{-\infty}^{+\infty} \overline{I(t)}^2 \cos 2\pi f\tau d\tau = \\ &= 2\nu \operatorname{Re} \left[\int_{-\infty}^{+\infty} dt F(t) e^{-2\pi i f t} \int_{-\infty}^{+\infty} dt' F(t') e^{2\pi i f t'} \right] + \\ &+ 2 \overline{I(t)}^2 \int_{-\infty}^{+\infty} e^{i2\pi f\tau} d\tau = 2\nu |s(f)|^2 + 2 \overline{I(t)}^2 \delta(f). \end{aligned} \quad (2.6-4)$$

При переходе от первого уравнения ко второму было принято $t' = t + \tau$ и $\cos 2\pi f\tau$ считался вещественной частью соответствующей показательной функции. При переходе от второго уравнения к третьему было положено

$$s(f) = \int_{-\infty}^{+\infty} F(t) e^{-2\pi i f t} dt, \quad (2.6-5)$$

а также использовано соотношение

$$\int_{-\infty}^{+\infty} e^{i2\pi f t} dt = \delta(f). \quad (2.2-9)$$

Второй член в $\overline{\omega(f)}$, имеющий в своем составе $\delta(f)$, характеризует среднюю мощность, которая выделялась бы постоянной составляющей $I(t)$, проходящей через сопротивление 1 ом. Это находится в соответствии с представлением о том, что средняя мощность в полосе частот $0 \leq f < \varepsilon$, где $\varepsilon > 0$, но очень мало, равна

$$\int_0^\varepsilon \overline{\omega(f)} df = 2 \overline{I(t)}^2 \int_0^\varepsilon \delta(f) df = \overline{I(t)}^2. \quad (2.6-6)$$

Уравнение (2.6—4) для $\bar{w}(f)$ может быть также получено из (2.5—3) путем прибавления дополнительного члена, связанного с постоянной составляющей, который получается усреднением выражения (2.2—11).

Интегрируя обе части (2.6—4) по f в пределах от 0 до ∞ и пользуясь соотношением

$$\bar{I}^2 = \int_0^{\infty} \bar{w}(f) df,$$

получим в результате

$$\bar{I}^2 - \bar{I}^2 = 2\nu \int_0^{\infty} |s(f)|^2 df. \quad (2.6—7)$$

Это уравнение может быть сразу получено из теоремы о наложении случайных возмущений путем применения теоремы Парсеваля.

В качестве примера использования этих формул найдем энергетический спектр напряжения на сопротивлении R , если ток состоит из большого числа очень коротких импульсов, протекающих через R . Пусть $F(t-t_k)$ будет напряжение, создаваемое импульсом, появившимся в момент t_k . Тогда

$$F(t) = R\varphi(t),$$

где $\varphi(t)$ есть ток в импульсе. Следует ограничиться сравнительно низкими частотами так, чтобы можно было воспользоваться приближением

$$s(f) = \int_{-\infty}^{+\infty} R\varphi(t)e^{-2\pi ift} dt \approx R \int_{-\infty}^{+\infty} \varphi(t) dt = Rq,$$

где q есть заряд, создаваемый одним импульсом. Из (2.6—4) следует, что на низких частотах сплошной участок энергетического спектра неизменен и равен

$$\bar{w}(f) = 2\nu R^2 q^2 = 2\bar{I}R^2 q, \quad (2.6—8)$$

где $\bar{I} = \nu q$ — средний ток, протекающий через R . Этот вывод часто используется в связи с дробовым эффектом в диодах.

При изучении дробового эффекта предполагалось, что вероятность того, что событие (поступление на анод электрона) произойдет в интервале dt , равна νdt , где ν есть ожидаемое число событий в 1 сек. Эта вероятность не зависит от времени t . Иногда желательно ввести зависимость от времени. В качестве примера рассмотрим длинный интервал, простирающийся от 0 до T . Пусть вероятность совершения события в промежутке $(t, t+dt)$ будет $\bar{K}p(t)dt$, где \bar{K} — среднее число событий за время T , а $p(t)$ есть такая данная

функция t , что

$$\int_0^T p(t) dt = 1$$

Для дробового эффекта $p(t) = \frac{1}{T}$.

Какова вероятность того, что точно K событий произойдет в интервале T ? Как и в случае дробового эффекта (раздел 1.1), можно разделить интервал $(0, T)$ на N промежутков каждый длительностью Δt так, что $N\Delta t = T$. Вероятность того, что ни одно событие не случится в первом промежутке Δt , равна

$$1 - \bar{K}p\left(\frac{\Delta t}{2}\right)\Delta t.$$

Произведение N таких вероятностей равно при $N \rightarrow \infty$, $\Delta t \rightarrow 0$

$$\exp\left[-\bar{K} \int_0^T p(t) dt\right] = e^{-\bar{K}}$$

Это вероятность того, что точно нуль событий произойдет за время T . Таким же образом приходим к выражению

$$\frac{\bar{K}^K}{K!} e^{-\bar{K}} \quad (2.6-9)$$

для вероятности того, что точно K событий случится за время T .

Рассмотрев много интервалов $(0, T)$, получим много значений K , а также много значений I , измеренных через t сек. от начала каждого интервала. Эти значения I определяют распределение I в момент t . Таким же образом, как и в разделе 1.4, найдем плотность вероятностей для I

$$P(I, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} du \exp\left\{-iul + \bar{K} \int_0^T p(x) [e^{-iuF(t-x)} - 1] dx\right\}$$

Соответствующие среднее значение и дисперсия равны

$$\begin{aligned} \bar{I} &= \bar{K} \int_0^T p(x) F(t-x) dx, \\ \overline{(I - \bar{I})^2} &= \bar{K} \int_0^T p(x) F^2(t-x) dx. \end{aligned} \quad (2.6-10)$$

Если $S(f)$ выражается уравнением (2.1-2), а $s(f)$ — (2.6-5), то, полагая длительность $F(t)$ малой сравнительно с T , среднее значение $|S(f)|^2$ можно получить, подставив (1.3-1) в (2.1-2):

$$S_K(f) = s(f) \sum_1^K e^{-2\pi i f t_k}$$

Представляя $S_K(f)S_K^*(f)$, где звездочка обозначает сопряженный комплекс, в виде двойной суммы и усредняя по t_k , используя выражение для $p(t)$ и затем усредняя по K , получим

$$\overline{[S(f)]^2} = \bar{K} |s(f)|^2 \left[1 + \bar{K} \int_0^T p(x) e^{-2\pi i f x} dx \right]^2 \quad (2.6-11)$$

Это соотношение может быть применено для нахождения энергетического спектра из (2.5—3), если только $p(x)$ не является периодической функцией. Если же $p(x)$ — периодическая функция, то тогда для гармонических составляющих должен быть применен метод раздела 2.2. Если флуктуации $p(t)$ медленны сравнительно с флуктуациями $F(t)$, то вторым членом внутри скобок в (2.6—11) обычно можно пренебречь, ибо нет таких значений f , при которых f и $s(f)$ были бы одновременно велики. Кроме того, если обе величины $p(t)$ и $F(t)$ подвержены флуктуациям примерно с одной и той же скоростью, то этот член должен быть учтен.

2.7. ВТОРОЙ ПРИМЕР — СЛУЧАЙНЫЙ ТЕЛЕГРАФНЫЙ СИГНАЛ

Пусть $I(t)$ равен либо a , либо $-a$, так что по форме ток представляет собой прямоугольную волну. Полагаем промежутки между переменами знака распределенными по показательному закону. Можно придти к такому распределению, считая, что если в среднем происходит μ перемен знака в 1 сек., то вероятность перемены знака в интервале $(t, t+dt)$ равна μdt и не зависит от того, что происходит вне интервала $(t, t+dt)$. Из рассуждений, аналогичных приведенным в разделе 1.1 для дробового эффекта, заключаем, что вероятность получения точно K перемен знака в интервале $(0, T)$ равна

$$p(K) = \frac{(\mu T)^K}{K!} e^{-\mu T} \quad (2.7-1)$$

Рассмотрим среднее значение произведения $I(t)I(t+\tau)$. Это произведение равно a^2 , если оба тока I одного знака, и $-a^2$, если они противоположного знака. В первом случае в интервале $(t, t+\tau)$ будет четное число, включая нуль, перемен знака, а во втором случае — нечетное число. Поэтому

$$\overline{I(t)I(t+\tau)} = a^2 \times [\text{вероятность четного числа перемен знака в промежутке } (t, t+\tau)] - a^2 \times [\text{вероятность нечетного числа перемен знака в промежутке } (t, t+\tau)]. \quad (2.7-2)$$

Длительность рассматриваемого интервала равна $|t+\tau-t| = |\tau|$ сек. Так как, по предположению, вероятность перемены знака в элементарном интервале Δt не зависит от происходящего за пределами этого интервала, то отсюда следует, что это предположение справедливо для любого интервала безотносительно от того,

когда он начинается. Следовательно, вероятности в (2.7—2) не зависят от t и могут быть определены из (2.7—1), полагая $T=|\tau|$. Тогда (2.7—2) превращается в следующее выражение, полагая $\tau > 0$:

$$\overline{I(t)I(t+\tau)} = a^2[p(0)+p(2)+p(4)+\dots] - a^2[p(1)+p(3)+ \\ + p(5)+\dots] = a^2 e^{-\mu\tau} \left[1 - \frac{\mu\tau}{1!} + \frac{(\mu\tau)^2}{2!} - \dots \right] = a^2 e^{-2\mu\tau} \quad (2.7-3)$$

Согласно (2.5—5), функция корреляции для $I(t)$ равна

$$\overline{\psi(\tau)} = a^2 e^{-2\mu|\tau|}. \quad (2.7-4)$$

Соответствующий энергетический спектр по (2.5—1) равен

$$\overline{\omega(f)} = 4a^2 \int_0^{\infty} e^{-2\mu\tau} \cos 2\pi f\tau \, d\tau = \frac{2a^2\mu}{\pi^2 f^2 + \mu^2} \quad (2.7-5)$$

Функции корреляции и энергетические спектры подобного типа встречаются весьма часто. В частности, они применяются при изучении турбулентности в гидродинамике. Можно также получить их и в случае дробового эффекта, если пренебречь постоянной составляющей. Все, что необходимо, это предположить, что эффект $F(t)$ поступления электрона на анод в момент $t=0$ при $t < 0$ равен нулю и что $F(t)$ спадает во времени по показательному закону после броска к максимальному значению в момент $t=0$. Это можно проверить, подставляя значение

$$F(t) = 2a \sqrt{\frac{\mu}{v}} e^{-2\mu t}, \quad t > 0 \quad (2.7-6)$$

в выражения (2.6—2) и (2.6—4) [после использования уравнения (2.6—5)] для функции корреляции и энергетического спектра дробового эффекта.

Энергетический спектр тока, протекающего через последовательное соединение из индуктивности и сопротивления под влиянием напряжения шумов теплового возбуждения в широкой полосе частот, также имеет вид (2.7—5).

Кстати, это дает пример двух совершенно различных токов $I(t)$, одного — прямоугольной формы, а второго — тока дробового эффекта, которые имеют одинаковые функции корреляции и энергетические спектры, если отвлечься от постоянной составляющей.

Есть еще другой тип случайного телеграфного сигнала, который интересно проанализировать. Ось времени делится на интервалы равной длительности h . В произвольно выбранном интервале значение $I(t)$ не зависит от значений в других интервалах и равно-

вероятно будет либо $+a$, либо $-a$. Можно имитировать такую волну, подбрасывая монету. Если выпадет герб, надо положить $I(t)=a$ в интервале $0 < t < h$. Если же выпадет решетка, то положим $I(t)=-a$ в этом же интервале. Вторичное подбрасывание монеты даст либо $+a$, либо $-a$ для второго интервала $h < t < 2h$ и т. д. Это дает одну волну. Подобным образом можно построить много волн. Усреднения по этим волнам при $t = \text{const}$ обозначены черточками.

Нас интересует среднее значение $I(t)I(t+\tau)$, полагая $\tau > 0$. Прежде всего замечаем, что если $\tau > h$, то эти два тока при любых значениях t будут относиться к различным интервалам. Так как значения в этих интервалах независимы друг от друга, то

$$\overline{I(t)I(t+\tau)} = \overline{I(t)} \overline{I(t+\tau)} = 0$$

для всех значений t , если $\tau > h$.

Чтобы найти среднее значение при $\tau < h$, рассмотрим случай, когда t заключено в первом интервале $0 < t < h$. Так как все интервалы со статистической точки зрения одинаковы, то общность при этом не теряется. Если $t+\tau < h$, т. е. $t < h-\tau$, то оба тока находятся в первом интервале и

$$\overline{I(t)I(t+\tau)} = a^2.$$

Если $t > h-\tau$, ток $I(t+\tau)$ находится во втором интервале, и тогда среднее значение равно нулю.

Возвратимся теперь к (2.5—4). Пределы интегрирования в этом уравнении от 0 до T . При $\tau > h$ подинтегральное выражение равно нулю, и тогда

$$\bar{\psi}(\tau) = 0, \quad \tau > h. \quad (2.7-7)$$

Если $\tau < h$, то исследования интервала $0 < t < h$ позволяют написать интеграл, соответствующий участку от 0 до h

$$\int_0^h \overline{I(t)I(t+\tau)} dt = \int_0^{h-\tau} a^2 dt + \int_{h-\tau}^h 0 dt = a^2(h-\tau)$$

По всему интервалу интегрирования $(0, T)$ будет T/h таких участков. Следовательно, из (2.5—4)

$$\bar{\psi}(\tau) = \lim_{T \rightarrow \infty} \frac{a^2}{T} \frac{T}{h} (h-\tau) = a^2 \left(1 - \frac{\tau}{h}\right) \quad 0 \leq \tau < h. \quad (2.7-8)$$

Энергетический спектр телеграфного сигнала такого типа поэтому равен

$$\bar{\omega}(f) = 4a^2 \int_0^h \left(1 - \frac{\tau}{h}\right) \cos 2\pi f\tau d\tau = 2h \left(\frac{a \sin \pi fh}{\pi fh}\right)^2 \quad (2.7-9)$$

Отсюда видно, что этот спектр имеет те же общие свойства, что и $\bar{\omega}(f)$ для первого типа телеграфного сигнала (2.7—5), если установить соотношение между μ — средним числом перемен знака в секунду и h — длительностью интервала в виде $\mu h = 1$.

2.8. ПРЕДСТАВЛЕНИЕ ТОКА ШУМОВ

В разделе 1.7 рассматривалось представление тока дробового эффекта в виде рядов Фурье. Это приводит к уравнению

$$I(t) = \sum_{n=1}^N (a_n \cos \omega_n t + b_n \sin \omega_n t), \quad (2.8-1)$$

где

$$\omega_n = 2\pi f_n, \quad f_n = n\Delta f. \quad (2.8-2)$$

Коэффициенты a_n и b_n принимаются за независимые случайные переменные, распределенные вокруг нуля по нормальному закону со стандартным отклонением $\sqrt{\omega(f_n)\Delta f}$. $\omega(f)$ есть энергетический спектр тока шумов, т.е. $\omega(f)\Delta f$ — средняя мощность, рассеиваемая теми составляющими $I(t)$, которые лежат в полосе частот $(f, f+df)$, если они протекают через сопротивление 1 ом.

Выражение для стандартного отклонения коэффициентов a_n и b_n получим, если заметим, что Δf есть ширина полосы частот, связанной с n -й составляющей. Тогда $\omega(f_n)\Delta f$ есть средняя мощность, которая выделялась бы, если бы ток

$$a_n \cos \omega_n t + b_n \sin \omega_n t$$

протекал через сопротивление 1 ом, причем это усреднение произведено по всем возможным значениям a_n и b_n . Поэтому

$$\begin{aligned} \omega(f_n)\Delta f &= \overline{a_n^2} \cos^2 \omega_n t + 2\overline{a_n b_n} \cos \omega_n t \sin \omega_n t + \\ &+ \overline{b_n^2} \sin^2 \omega_n t = \overline{a_n^2} = \overline{b_n^2}. \end{aligned} \quad (2.8-3)$$

Последнее следует из независимости a_n и b_n друг от друга и идентичности их распределений. Можно заметить, что $\omega(f)$, полученное в связи с представлением (2.8—1), есть среднее того же типа, что и обозначенное в разделе 2.5 через $\bar{\omega}(f)$.

Например, допустим, что нас интересует ток на выходе некоторого фильтра, когда на входе действует источник термических шумов. Пусть $A(f)$ будет абсолютное значение отношения выходного тока ко входному, когда на вход подведено установившееся синусоидальное напряжение частоты f . Тогда

$$\omega(f) = \zeta A^2(f). \quad (2.8-4)$$

Пусть W есть средняя мощность, рассеиваемая на сопротивлении 1 ом током $I(t)$:

$$W = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T I^2(t) dt = \int_0^{\infty} \omega(f) df = \zeta \int_0^{\infty} A^2(f) df; \quad (2.8-5)$$

из этого уравнения можно найти ζ , если известны W и $A(f)$.

Пользуясь (2.8—1) для исследования статистических свойств $I(t)$, сначала находим соответствующие статистические свойства суммы, стоящей справа, в которой a и b рассматриваются как случайные переменные, распределенные, как указывалось выше, а t полагается фиксированным. Вообще же время t исчезает при этой процедуре, так же как и в (2.8—3). Положим, далее, $N \rightarrow \infty$, а $\Delta f \rightarrow 0$, так что суммирование можно заменить интегрированием, и наконец, расширим диапазон частот, чтобы перекрыть все частоты от 0 до ∞ .

Обычный путь использования уравнения (2.8—1) заключается в предположении, что имеется осциллограмма $I(t)$, простирающаяся от $t=0$ до $t=\infty$. Эта осциллограмма может быть разрезана на полосы длиной T . Разложение тока $I(t)$ каждой полосы в ряд Фурье (при $T\Delta f=1$) даст группу коэффициентов, которые будут изменяться от полосы к полоске. Предполагаем, что это изменение подчиняется нормальному закону распределения.

Применяемый здесь процесс нахождения статистических свойств путем рассмотрения a и b как случайных переменных при t , сохраняемом постоянным, соответствует исследованию значений тока шумов в большом числе моментов времени. Каждой полоске соответствует определенный момент, и этот момент выбирается через t секунд от начала полосы. [Это есть t , входящее в (2.8—1)]. До некоторой степени такой подход похож на исследование тока шумов в большом числе случайно выбранных моментов.

Хотя уравнение (2.8—1) может быть использовано для представления дробового эффекта и других подобных явлений, оно не является единственным, а в ряде случаев не оказывается и наиболее удобным. Другое представление, приводящее к тем же самым выводам:

$$I(t) = \sum_{n=1}^N c_n \cos(\omega_n t - \varphi_n), \quad (2.8-6)$$

где $\varphi_1, \varphi_2, \dots, \varphi_N$ — углы, случайно распределенные в интервале $(0, 2\pi)$, а

$$c_n = \sqrt{2\omega(f_n)\Delta f}, \quad \omega_n = 2\pi f_n, \quad f_n = n\Delta f. \quad (2.8-7)$$

При таком представлении $I(t)$ рассматривается как сумма синусоидальных составляющих с постоянными амплитудами, но со случайными фазовыми углами.

То обстоятельство, что два разных представления (2.8—1) и (2.8—6) приводят к одинаковым статистическим свойствам, есть следствие того, что в обоих случаях может быть использована центральная предельная теорема¹⁾.

Эта теорема утверждает, что при некоторых общих условиях распределение суммы N случайных векторов сходится к нормальному закону (закон может быть нормальным в нескольких измерениях²⁾) при $N \rightarrow \infty$. Действительно, из этой теоремы следует, что представление тока в виде

$$I(t) = \sum_{n=1}^N (a_n \cos \omega_n t + b_n \sin \omega_n t), \quad (2.8-1)$$

где a_n и b_n — независимые случайные переменные, принимающие только значения $\pm \sqrt{\omega(f_n)\Delta f}$, причем вероятность каждого значения равна $\frac{1}{2}$, приводит в пределе к таким же точно статистическим свойствам, что и (2.8—6).

2.9. НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

Рассмотрим случайный вектор r , находящийся в пространстве K измерений.

Распределение этого вектора может быть определено указанием распределения K составляющих x_1, x_2, \dots, x_K вектора r . Говорят, что r распределен по нормальному закону, когда функция плотности вероятностей для r имеет вид

$$(2\pi)^{-\frac{K}{2}} |M|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} x' M^{-1} x\right) \quad (2.9-1)$$

где экспонента есть квадратичная форма от x . Квадратная матрица M составлена из моментов второго порядка

$$M = \begin{vmatrix} \mu_{11} & \mu_{12} & \dots & \mu_{1K} \\ \dots & \dots & \dots & \dots \\ \mu_{1K} & \dots & \dots & \mu_{KK} \end{vmatrix} \quad (2.9-2)$$

где моменты определяются как

$$\mu_{11} = \overline{x_1^2}, \quad \mu_{12} = \overline{x_1 x_2}, \quad \text{и т. д.} \quad (2.9-3)$$

$|M|$ есть определитель матрицы M , а x' — матрица строк

$$x' = [x_1, x_2, \dots, x_K]. \quad (2.9-4)$$

¹⁾ См. 2.10.

²⁾ См. 2.9.

x — матрица столбцов, полученная путем транспонирования x' Экспонента в выражении (2.9—1) для плотности вероятностей может быть выписана полностью при использовании соотношения

$$x' M^{-1} x = \sum_{r=1}^K \sum_{s=1}^K \frac{M_{rs}}{|M|} x_r x_s, \quad (2.9-5)$$

где M_{rs} — алгебраическое дополнение элемента μ_{rs} в матрице M .

Иногда между составляющими x имеются линейные зависимости, так что случайный вектор r ограничен пространством меньше, чем K измерений. В этом случае подходящая форма функции плотности вероятностей может быть получена из последовательности K -мерных распределений, сходящихся к только что рассмотренному.

Если r_1 и r_2 — два нормально распределенных случайных вектора, то их сумма $r_1 + r_2$ также распределена нормально. Отсюда следует, что сумма любого числа нормально распределенных векторов распределена по нормальному закону.

Характеристическая функция нормального распределения равна

$$\begin{aligned} \text{сред.}[\exp(i z_1 x_1 + i z_2 x_2 + \dots + i z_K x_K)] &= \\ &= \exp \left[-\frac{1}{2} \sum_{r=1}^K \sum_{s=1}^K \mu_{rs} z_r z_s \right] \end{aligned} \quad (2.9-6)$$

2.10. ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА

Центральная предельная теорема теории вероятностей утверждает, что распределение суммы из N независимых случайных векторов $r_1 + r_2 + \dots + r_N$ сходится к нормальному закону при $N \rightarrow \infty$, если распределения векторов r_1, r_2, \dots, r_N удовлетворяют некоторым общим условиям.

В качестве примера возьмем случай, когда r_1, r_2, \dots, r_N — векторы двух измерений, причем составляющими r_n являются x_n и y_n . Не теряя общности, допустим, что

$$\bar{x}_n = 0, \quad \bar{y}_n = 0.$$

Составляющие результирующего вектора равны

$$\begin{aligned} X &= x_1 + x_2 + \dots + x_N, \\ Y &= y_1 + y_2 + \dots + y_N, \end{aligned} \quad (2.10-1)$$

а так как r_1, r_2, \dots — независимые векторы, то моменты второго порядка результирующего вектора равны

$$\begin{aligned} \mu_{11} &= \overline{X^2} = \overline{x_1^2} + \overline{x_2^2} + \dots + \overline{x_N^2}, \\ \mu_{22} &= \overline{Y^2} = \overline{y_1^2} + \overline{y_2^2} + \dots + \overline{y_N^2}, \\ \mu_{12} &= \overline{XY} = \overline{x_1 y_1} + \overline{x_2 y_2} + \dots + \overline{x_N y_N}. \end{aligned} \quad (2.10-2)$$

Очевидно, существует несколько различных условий, выполнение которых достаточно для того, чтобы распределение результирующего вектора сходилось к нормальному закону. Одно из достаточных условий

$$\begin{aligned} \mu_{11}^{-3/2} \sum_{n=1}^N |x_n|^3 &\rightarrow 0, \\ \mu_{22}^{-3/2} \sum_{n=1}^N |y_n|^3 &\rightarrow 0. \end{aligned} \quad (2.10-3)$$

Центральная предельная теорема утверждает, что распределение случайного вектора (X, Y) сходится к нормальному закону при $N \rightarrow \infty$. Моменты второго порядка этого распределения даются уравнением (2.10-2). Если известны моменты второго порядка нормального распределения, то можно сразу написать функцию плотности вероятностей. Поэтому из раздела 2.9

$$\begin{aligned} M &= \begin{vmatrix} \mu_{11} & \mu_{12} \\ \mu_{12} & \mu_{22} \end{vmatrix}, & M^{-1} &= |M|^{-1} \begin{vmatrix} \mu_{22} & -\mu_{12} \\ -\mu_{12} & \mu_{11} \end{vmatrix}, \\ |M| &= \mu_{11}\mu_{22} - \mu_{12}^2, \\ x' &= [X, Y], \\ x' M^{-1} x &= |M|^{-1} (\mu_{22} X^2 - 2\mu_{12} XY + \mu_{11} Y^2). \end{aligned}$$

Следовательно, плотность вероятностей равна

$$\frac{(\mu_{11}\mu_{22} - \mu_{12}^2)^{-1/2}}{2\pi} \exp \left[\frac{-\mu_{22} X^2 - \mu_{11} Y^2 + 2\mu_{12} XY}{2(\mu_{11}\mu_{22} - \mu_{12}^2)} \right] \quad (2.10-4)$$

Моменты второго порядка связаны со стандартными отклонениями для X , Y , обозначаемыми σ_1 и σ_2 , и с коэффициентом корреляции для X и Y , обозначаемым τ , следующим образом:

$$\mu_{11} = \sigma_1^2, \quad \mu_{22} = \sigma_2^2, \quad \mu_{12} = \tau \sigma_1 \sigma_2, \quad (2.10-5)$$

и тогда плотность вероятностей принимает обычную форму

$$\frac{(1-\tau^2)^{-1/2}}{2\pi\sigma_1\sigma_2} \exp \left[-\frac{1}{2(1-\tau^2)} \left(\frac{X^2}{\sigma_1^2} - 2\tau \frac{XY}{\sigma_1\sigma_2} + \frac{Y^2}{\sigma_2^2} \right) \right]. \quad (2.10-6)$$

Г л а в а III

СТАТИСТИЧЕСКИЕ СВОЙСТВА ФЛУКТУАЦИОННЫХ ШУМОВ

В этой главе применены изложенные в разделе 2.8 представления о шумовых токах для вывода некоторых статистических свойств $I(t)$. Первые шесть разделов посвящены распределению вероятностей тока $I(t)$ и его нулям и выбросам. Разделы 3.7 и 3.8 связаны со статистическими свойствами огибающей $I(t)$. В разделе 3.9 рассмотрены флуктуации интегралов, в которые входит $I^2(t)$. Распределению вероятностей суммы из синусоидального тока и тока шумов посвящен раздел 3.10, а в разделе 3.11 кратко описан другой метод получения выводов главы III. Большинство материала этой главы тесно связано с теорией процессов Маркова.

3.1. РАСПРЕДЕЛЕНИЕ ТОКА ШУМОВ

В разделе 1.4 было показано, что распределение тока дробового эффекта сходится к нормальному закону, когда ожидаемое число событий в 1 сек. ν безгранично возрастает.

В соответствии с задачами этой части будем пользоваться

$$I(t) = \sum_{n=1}^N (a_n \cos \omega_n t + b_n \sin \omega_n t) \quad (2.8-1)$$

для того чтобы показать, что функция $I(t)$ распределена по нормальному закону. Этот результат немедленно может быть получен, если следовать процедуре раздела 2.8. Так как a_n и b_n распределены нормально, то таково же распределение $a_n \cos \omega_n t$ и $b_n \sin \omega_n t$, если рассматривать t как фиксированное. Поэтому $I(t)$ есть сумма $2N$ независимых нормально распределенных переменных, а следовательно, и сама распределена по нормальному закону.

Среднее значение $I(t)$ по (2.8—1) равно нулю, так как

$$\begin{aligned} \overline{a_n} = \overline{b_n} &= 0, \\ \overline{I(t)} &= 0. \end{aligned} \quad (3.1-1)$$

Средний квадрат $I(t)$ равен

$$\overline{I^2(t)} = \sum_{n=1}^N (\overline{a_n^2} \cos^2 \omega_n t + \overline{b_n^2} \sin^2 \omega_n t) =$$

$$\begin{aligned}
 &= \sum_{n=1}^N \omega(f_n) \Delta f \rightarrow \\
 &\rightarrow \int_0^{\infty} \omega(f) df = \psi(0) \equiv \psi_0.
 \end{aligned}
 \tag{3.1-2}$$

В верхней строчке (3.1—2) было использовано то обстоятельство, что все a и b независимы и, следовательно, среднее любого комбинированного произведения равно нулю. Были также применены результаты раздела 2.8

$$\bar{a}_n^2 = \bar{b}_n^2 = \omega(f_n) \Delta f, \quad f_n = n \Delta f, \quad \omega_n = 2\pi f_n.$$

Как показано в разделе 2.1, $\psi(\tau)$ есть функция корреляции для $I(t)$, связанная с $\omega(f)$ следующим образом:

$$\psi_\tau \equiv \psi(\tau) = \int_0^{\infty} \omega(f) \cos 2\pi f \tau df.
 \tag{2.1-6}$$

В этой главе для краткости аргумент $\psi(\tau)$ записан в виде индекса.

Так как известно, что $I(t)$ распределен по нормальному закону, его среднее значение равно нулю, а средний квадрат есть ψ_0 , то можно сразу же написать и функцию плотности вероятностей. Поэтому вероятность того, что $I(t)$ находится в интервале $(I, I + dI)$, равна

$$\frac{dI}{\sqrt{2\pi\psi_0}} e^{-I^2/2\psi_0}
 \tag{3.1-3}$$

Это есть вероятность нахождения тока в интервале между I и $I + dI$ в случайно выбранный момент времени. Другими словами, уравнение (3.1—3) дает ту долю времени, в течение которой ток находится в интервале $(I, I + dI)$.

Во многих случаях более удобно пользоваться выражением

$$I(t) = \sum_{n=1}^N c_n \cos(\omega_n t - \varphi_n), \quad c_n^2 = 2\omega(f_n) \Delta f,
 \tag{2.8-6}$$

где $\varphi_1, \dots, \varphi_N$ — независимые случайные фазовые углы. Чтобы показать, что и в этом виде $I(t)$ подчиняется нормальному распределению, заметим прежде всего, что в (2.8—6) $I(t)$ представлен в виде суммы большого числа независимых случайных переменных

$$\begin{aligned}
 I(t) &= x_1 + x_2 + \dots + x_N, \\
 x_n &= c_n \cos(\omega_n t - \varphi_n),
 \end{aligned}$$

и, следовательно, когда $N \rightarrow \infty$, то $I(t)$ будет распределен по нормальному закону. Чтобы сделать предельный процесс конечным, вначале выберем N и Δf такими, что $N\Delta f = F$, где

$$\int_F \omega(f) df < \varepsilon \int_0^\infty \omega(f) df,$$

а ε — некоторая произвольно выбранная малая положительная величина. Положим теперь $N \rightarrow \infty$, а $\Delta f \rightarrow 0$ таким образом, что $N\Delta f$ остается равным F . Тогда

$$A = \overline{x_1^2} + \overline{x_2^2} + \dots + \overline{x_N^2} = \sum_1^N 2\omega(f_n)\Delta f \overline{\cos^2(\omega_n t - \varphi_n)} = \\ = \sum_1^N \omega(f_n)\Delta f \rightarrow \int_0^F \omega(f) df,$$

$$B = \overline{|x_1|^3} + \dots + \overline{|x_N|^3} = \sum_1^N [2\omega(f_n)\Delta f]^{3/2} \overline{|\cos(\omega_n t - \varphi_n)|^3} < \\ < 4(\Delta f)^{3/2} \int_0^F [\omega(f)]^{3/2} df, \quad (3.1-4)$$

где черта над членами обозначает усреднение по φ , а t поддерживается постоянным. Если предположить, что интегралы не являются расходящимися, то отношение $BA^{-3/2} \rightarrow 0$, когда $N \rightarrow \infty$. Следовательно, можно воспользоваться центральной предельной теоремой¹⁾, если $\omega(f) = 0$ при $f > F$. Так как F может быть взято сколь угодно большим путем выбора ε достаточно малым, то можно перекрыть сколь угодно большой диапазон частот. По этой причине вместо F записывается ∞ .

Теперь, когда при помощи центральной предельной теоремы было показано, что распределение для $I(t)$ в виде (2.8—6) сходится к нормальному закону, остается только найти среднее значение и средний квадрат для $I(t)$:

$$\overline{I(t)} = \sum_1^N c_n \overline{\cos(\omega_n t - \varphi_n)} = 0, \\ \overline{I^2(t)} = \sum_1^N c_n^2 \overline{\cos^2(\omega_n t - \varphi_n)} \rightarrow \int_0^\infty \omega(f) df = \psi_0 \quad (3.1-5)$$

Отсюда получаем плотность вероятностей согласно (3.1—3); таким образом, оба представления в этом случае приводят к одинаковым результатам. Очевидно, одинаковые результаты будут получаться до тех пор, пока можно применять центральную предель-

¹⁾ Раздел 2.10.

ную теорему. В дальнейшем, пользуясь (2.8—6), просто будем предполагать, что для доказательства сходимости к нормальному распределению может быть использована центральная предельная теорема, а все вычисления, относящиеся к уравнениям (3.1—4), будут опускаться.

Характеристическая функция для распределения $I(t)$ равна

$$\text{сред. } e^{iuI(t)} = \exp\left(-\frac{\psi_0}{2}u^2\right) \quad (3.1-6)$$

3.2. РАСПРЕДЕЛЕНИЕ $I(t)$ и $I(t + \tau)$

Нужно найти распределение двух величин, в котором первая переменная есть ток шумов $I(t)$, а вторая переменная — значение этого тока $I(t + \tau)$ через некоторый промежуток времени τ . Оказывается, что это распределение нормально, как и можно было ожидать по аналогии с разделом 3.1. Моменты второго порядка этого распределения равны

$$\begin{aligned} \mu_{11} &= \overline{I^2(t)} = \psi_0 = \int_0^\infty \omega(f) df, \\ \mu_{22} &= \psi_0, \\ \mu_{12} &= \overline{I(t)I(t + \tau)} = \psi_\tau. \end{aligned} \quad (3.2-1)$$

Выражение для μ_{12} соответствует нашему определению (2.1—4) функции корреляции

$$\psi_\tau \equiv \psi(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T I(t)I(t + \tau) dt. \quad (2.1-4)$$

Чтобы найти распределение, пользуясь выражением (2.8—6), напомним

$$\begin{aligned} I_1 &= I(t) = \sum_1^N c_n \cos(\omega_n t - \varphi_n), \\ I_2 &= I(t + \tau) = \sum_1^N c_n \cos(\omega_n t - \varphi_n + \omega_n \tau). \end{aligned}$$

Из центральной предельной теоремы для двух измерений следует, что I_1 и I_2 распределены нормально. Так же как и в разделе (3.1):

$$\begin{aligned} \mu_{11} &= \overline{I_1^2} = \sum_1^N c_n^2 \frac{1}{2} \rightarrow \int_0^\infty \omega(f) df = \psi_0, \\ \mu_{22} &= \overline{I_2^2} = \overline{I_1^2} = \psi_0, \\ \mu_{12} &= \overline{I_1 I_2} = \sum_1^N c_n^2 [\cos(\omega_n t - \varphi_n) \cos(\omega_n t - \varphi_n + \omega_n \tau)]. \end{aligned} \quad (3.2-2)$$

Величина, помещенная в квадратные скобки, равна

$$\cos^2(\omega_n t - \varphi_n) \cos \omega_n \tau - \cos(\omega_n t - \varphi_n) \sin(\omega_n t - \varphi_n) \sin \omega_n \tau,$$

и если взять усреднение по φ_n , то второй член пропадает, и тогда

$$\mu_{12} = \sum_1^N c_n^2 \frac{1}{2} \cos \omega_n \tau \rightarrow \int_0^\infty \omega(f) \cos 2\pi f \tau df = \psi_\tau, \quad (3.2-3)$$

где использовано соотношение (2.1—6) между $\omega(f)$ и $\psi(\tau)$ и применено обозначение $\omega_n = 2\pi f_n$.

Теперь может быть написана функция плотности вероятностей для I_1 и I_2 . Из рассмотрения нормального закона в разделе 2.9 следует, что она равна

$$\frac{(\psi_0^2 - \psi_\tau^2)^{-1/2}}{2\pi} \exp \left[\frac{-\psi_0 I_1^2 - \psi_0 I_2^2 + 2\psi_\tau I_1 I_2}{2(\psi_0^2 - \psi_\tau^2)} \right] \quad (3.2-4)$$

Для полосового фильтра с полосой пропускания от f_a до f_b имеем

$$\psi_\tau = \int_{f_a}^{f_b} \omega_0 \cos 2\pi f \tau df = \omega_0 \frac{\sin \omega_b \tau - \sin \omega_a \tau}{2\pi\tau} =$$

$$= \frac{\omega_0}{\pi\tau} \sin \pi\tau (f_b - f_a) \cos \pi\tau (f_b + f_a), \quad (3.2-5)$$

$$\psi_0 = \omega_0 (f_b - f_a),$$

где ω_0 есть постоянное значение $\omega(f)$ в полосе пропускания, а

$$\omega_b = 2\pi f_b, \quad \omega_a = 2\pi f_a. \quad (3.2-6)$$

Согласно (3.2—4) I_1 и I_2 независимы, когда $\psi_\tau = 0$. Для тех значений τ , при которых $\psi_\tau = 0$, знание I_1 не увеличивает наши сведения об I_2 . Например, допустим, что имеется узкополосный фильтр. Тогда

$$\psi_\tau = 0, \text{ когда } \tau = [2(f_b - f_a)]^{-1}$$

$$\psi_\tau \text{ почти равно } -\psi_0, \text{ когда } \tau = (f_b + f_a)^{-1}.$$

При первом значении τ известно только, что I_2 распределен около нуля с $\bar{I}_2^2 = \psi_0$. При втором значении τ I_2 , вероятно, почти равен $-I_1$. Это согласуется с представлением о том, что ток шумов, прошедший через узкополосный фильтр, ведет себя подобно

синусоидальной волне с частотой $\frac{1}{2}(f_b + f_a)$, а что касается его амплитуды, то она подвергается флуктуациям с нерегулярной частотой порядка $\frac{1}{2}(f_b - f_a)$. Первое значение τ соответствует четверти периода подобной волны, а второе — половине периода. Начертив синусоидальную волну и рассмотрев точки, разделенные промежутками в четверть периода и в полпериода, можно убедиться, что подобное представление подтверждается.

Характеристическая функция распределения I_1 и I_2 равна

$$\text{сред. } e^{iuI_1 + ivI_2} = \exp \left[-\frac{\psi_0}{2} (u^2 + v^2) - \psi_\tau uv \right]. \quad (3.2-7)$$

Распределение трех величин

$$I_1 = I(t), \quad I_2 = I(t + \tau_1), \quad I_3 = I(t + \tau_1 + \tau_2),$$

где τ_1 и τ_2 заданы, а t выбрано случайным, является, как и следовало ожидать, нормальным во всех трех измерениях. Моменты, из которых по методу раздела 2.9 можно найти распределение, равны

$$\mu_{11} = \mu_{22} = \mu_{33} = \psi_0,$$

$$\mu_{12} = \psi_{\tau_1},$$

$$\mu_{23} = \psi_{\tau_2},$$

$$\mu_{13} = \psi(\tau_1 + \tau_2) = \psi_{\tau_1 + \tau_2}.$$

Характеристическая функция для I_1 , I_2 и I_3 равна

$$\text{сред. } e^{iz_1 I_1 + iz_2 I_2 + iz_3 I_3} =$$

$$= \exp \left[-\frac{\psi_0}{2} (z_1^2 + z_2^2 + z_3^2) - \mu_{12} z_1 z_2 - \mu_{23} z_2 z_3 - \mu_{13} z_1 z_3 \right]. \quad (3.2-8)$$

3.3. ОЖИДАЕМОЕ ЧИСЛО НУЛЕЙ В 1 СЕК.¹⁾

Пусть y равно

$$y = F(a_1, a_2, \dots, a_N; x) \quad (3.3-1)$$

и пусть a — случайные переменные. Для данного ряда значений a это уравнение дает кривую зависимости y от x . Так как a являются случайными переменными, то назовем эту кривую случайной. Возьмем короткий интервал $(x_1, x_1 + dx)$ и затем выберем некоторый

¹⁾ Под этим понимается ожидаемое число проходов кривой тока шумов через уровень, относительно которого распределение симметрично. (Прим. ред.)

ряд значений a . Вероятность того, что кривая, полученная в результате подстановки этих a в уравнение (3.3—1), будет иметь нуль в интервале $(x_1, x_1 + dx)$, равна

$$dx \int_{-\infty}^{+\infty} |\eta| p(0, \eta; x_1) d\eta. \quad (3.3-2)$$

Ожидаемое число нулей в интервале (x_1, x_2) равно

$$\int_{x_1}^{x_2} dx \int_{-\infty}^{+\infty} |\eta| p(0, \eta; x) d\eta. \quad (3.3-3)$$

В этих выражениях $p(\xi, \eta; x)$ есть функция плотности вероятностей для переменных

$$\xi = F(a_1, \dots, a_N; x), \quad \eta = \frac{\partial F}{\partial x}. \quad (3.3-4)$$

Так как a — случайные переменные, то таковы же ξ и η и в их распределение будет входить x как параметр, что указывается обозначением $p(\xi, \eta; x)$.

Эти выводы могут быть получены таким же образом, как это сделано для случая распределения выбросов случайной кривой. Этому методу доказательства свойствен тот недостаток, что в нем требуется, чтобы a были ограничены.

Теперь перейдем к доказательству близко связанного с предыдущим положения: вероятность прохождения y через нуль в интервале $(x_1, x_1 + dx)$ с положительной крутизной равна

$$dx \int_0^{\infty} \eta p(0, \eta; x_1) d\eta. \quad (3.3-5)$$

dx выбирается настолько малым, что участки почти всех возможных случайных кривых (за исключением ничтожной части), относящиеся к интервалу $(x_1, x_1 + dx)$, могут рассматриваться как прямые линии. Если $y = \xi$ при x_1 и проходит через нуль при $x_1 < x < x_1 + dx$, то отрезок, отсекаемый на оси x при $y = 0$, равен $x_1 - \frac{\xi}{\eta}$, где η — крутизна. Поэтому ξ и η должны иметь противоположные знаки и

$$x_1 < x_1 - \frac{\xi}{\eta} < x_1 + dx.$$

Согласно формулировке задачи нас интересуют только положительные значения η , и поэтому напомним это неравенство в виде

$$-\eta dx < \xi < 0.$$

Для данной случайной кривой, т. е. для данного ряда a , значения ξ и η равны

$$\xi = F(a_1, \dots, a_N; x_1), \quad \eta = \left[\frac{\partial F}{\partial x} \right]_{x=x_1}.$$

Если эти значения ξ и η удовлетворяют нашему неравенству, то кривая проходит через нуль в интервале $(x_1, x_1 + dx)$. Вероятность этого равна ¹⁾

$$\int_0^{\infty} d\eta \int_{-\eta dx}^0 d\xi p(\xi, \eta; x_1) = \int_0^{\infty} [0 - (-\eta dx)] p(0, \eta; x_1) d\eta,$$

где учтено, что поскольку dx весьма мало, то ξ равно нулю. Последнее выражение совпадает с (3.3—5).

Таким же образом можно показать, что вероятность прохождения y через нуль в интервале $(x_1, x_1 + dx)$ с отрицательной крутизной равна

$$-dx \int_{-\infty}^0 \eta p(0, \eta; x_1) d\eta. \quad (3.3-6)$$

Выражение (3.3—2) получается путем сложения (3.3—5) и (3.3—6).

Далее можно перейти к применению выведенных формул. Заменяем x , y и a_n на t , $I(t)$ и φ_n , соответственно, и воспользуемся соотношениями

$$I(t) = \sum_{n=1}^N c_n \cos(\omega_n t - \varphi_n), \quad c_n^2 = 2\omega(f)\Delta f. \quad (2.8-6)$$

¹⁾ Переход от двойного интеграла в левой части этого уравнения к конечному результату (3.3—5) можно выполнить следующим образом: легко видеть, что искомая плотность вероятностей равна

$$\left[\frac{d}{d(\Delta x)} \int_0^{\infty} d\eta \int_{-\eta \Delta x}^0 p(\xi, \eta; x) d\xi \right]_{\Delta x=0}.$$

Действуя формально, безотносительно от выполнения условий, оправдывающих аналитические преобразования (они выполняются в интересующих нас случаях), получим

$$\frac{d}{d\Delta x} \int_0^{\infty} d\eta \int_{-\eta \Delta x}^0 p(\xi, \eta; x) d\xi = \int_0^{\infty} \eta p(-\eta \Delta x, \eta; x) d\eta,$$

а отсюда искомая плотность вероятностей равна

$$\int_0^{\infty} \eta p(0, \eta; x) d\eta.$$

Первым шагом является нахождение функции плотности вероятностей двух случайных переменных

$$\xi = \sum_{n=1}^N c_n \cos(\omega_n t_1 - \varphi_n), \quad (3.3-7)$$

$$\eta = I'(t_1) = - \sum_{n=1}^N c_n \omega_n \sin(\omega_n t_1 - \varphi_n),$$

где штрих обозначает дифференцирование по t . Из раздела 2.10

$$\mu_{11} = \overline{\xi^2} = \psi_0,$$

$$\begin{aligned} \mu_{22} = \overline{\eta^2} &= \sum_{n=1}^N c_n^2 \omega_n^2 \overline{\sin^2(\omega_n t_1 - \varphi_n)} = \sum_{n=1}^N (2\pi f_n)^2 \omega(f_n) \Delta f \rightarrow \\ &\rightarrow 4\pi^2 \int_0^{\infty} f^2 \omega(f) df = -\psi_0'' \end{aligned}$$

$$\mu_{12} = \overline{\xi\eta} = - \sum_{n=1}^N c_n^2 \omega_n \overline{\cos(\omega_n t_1 - \varphi_n) \sin(\omega_n t_1 - \varphi_n)} = 0.$$

Выражение для μ_{22} получается из (2.1—6) путем дифференцирования. В этом выражении ψ_0'' обозначает вторую производную $\psi(\tau)$ по τ при $\tau = 0$

$$\psi''(\tau) = -4\pi^2 \int_0^{\infty} f^2 \omega(f) \cos 2\pi f \tau df. \quad (3.3-8)$$

Отсюда плотность вероятностей равна

$$p(\xi, \eta; t) = \frac{(-\psi_0 \psi_0'')^{-1/2}}{2\pi} \exp\left(-\frac{\xi^2}{2\psi_0} + \frac{\eta^2}{2\psi_0''}\right), \quad (3.3-9)$$

где ψ_0'' отрицательно. Заметим, что выражение справа не зависит от t . Отсюда вероятность нахождения нуля в интервале $(t_1, t_1 + dt)$

$$dt \int_{-\infty}^{+\infty} |\eta| \frac{(-\psi_0 \psi_0'')^{-1/2}}{2\pi} e^{\eta^2/2\psi_0''} d\eta = \frac{dt}{\pi} \left[-\frac{\psi''(0)}{\psi(0)} \right]^{1/2}, \quad (3.3-10)$$

получающаяся из (3.3—2), не зависит от t .

Ожидаемое число нулей в 1 сек., которое можно найти из (3.3—3) путем интегрирования (3.3—10) по интервалу в 1 сек., равно

$$\frac{1}{\pi} \left[\frac{-\psi''(0)}{\psi(0)} \right]^{1/2} = 2 \left[\frac{\int_0^\infty f^2 \omega(f) df}{\int_0^\infty \omega(f) df} \right]^{1/2} \quad (3.3-11)$$

Для идеального полосового фильтра, полоса пропускания которого простирается от f_a до f_b , ожидаемое число нулей в 1 сек. равно

$$2 \left[\frac{1}{3} \frac{f_b^3 - f_a^3}{f_b - f_a} \right]^{1/2} \quad (3.3-12)$$

Когда $f_a=0$, это выражение равно $1,155 f_b$, а когда f_a весьма близко к f_b , то оно стремится к $f_b + f_a$.

В недавней работе М. Кэк¹⁾ дает выражение, которое после небольшого обобщения приводит к

$$e^{-I^2/2\psi_0} \frac{1}{2\pi} \left(\frac{-\psi_0''}{\psi_0} \right)^{1/2} dt \quad (3.3-13)$$

для вероятности того, что ток шумов пройдет через значение I с положительной крутизной в интервале $(t, t+dt)$. Ожидаемое число таких прохождений в 1 сек. равно

$$e^{-I^2/2\psi_0} \times \left[\frac{1}{2} \text{ожидаемое число нулей в 1 сек.} \right]. \quad (3.3-14)$$

Уравнение (3.3-13) может быть также получено из выражения, аналогичного (3.3-5), в котором в $p(0, \eta; x_1)$ нуль заменен на y .

В некоторых случаях интеграл

$$\psi_0'' = -4\pi^2 \int_0^\infty f^2 \omega(f) df$$

не сходится.

В качестве примера можно указать на случай воздействия на цепь, состоящую из последовательно включенных сопротивления и конденсатора, напряжения шумов, занимающего широкую полосу частот. Энергетический спектр напряжения на конденсаторе имеет вид

$$\omega(f) = \frac{1}{f^2 + a^2}. \quad (3.3-15)$$

Хотя ψ_0'' бесконечно велико, ψ_0 конечно и равно $\pi/2a$. Непосредственная подстановка в (3.3-11) дает для ожидаемого числа нулей в 1 сек. бесконечное значение.

¹⁾ См. «О распределении значений тригонометрических сумм с линейно независимыми частотами», *Amer. Journ. Math.*, LXV, 609—615, 1943.

Формула (3.3.—11) станет яснее при рассмотрении тока шумов, спектр которого состоит из двух частотных полос. Одна полоса ограничена сравнительно низкими частотами и ее энергетический спектр обозначим $\omega_1(f)$. Другая полоса весьма узка и ее центральной частотой является сравнительно высокая частота f_2 . Тогда полный энергетический спектр шумов

$$\omega(f) = \omega_1(f) + A^2 \delta(f - f_2),$$

где единичная импульсная функция δ используется для представления весьма узкой полосы. Энергетический спектр узкой полосы примерно одинаков со спектром волны, уравнение которой $A\sqrt{2} \cos 2\pi f_2 t$.

Интегралы, встречающиеся в формулах, имеют вид

$$\int_0^{\infty} \omega(f) df = \int_0^{\infty} \omega_1(f) df + A^2 = W + A^2,$$

$$\int_0^{\infty} \omega(f) f^2 df = \int_0^{\infty} f^2 \omega_1(f) df + A^2 f_2^2 = U + A^2 f_2^2.$$

Пусть A и f_2 таковы, что

$$W \gg A^2, \quad U \ll A^2 f_2^2.$$

Тогда формула (3.3.—11) дает ожидаемое число нулей

$$2 \frac{A f_2}{W^{1/2}}.$$

Можно дать качественное объяснение этой формуле, если рассматривать ток шумов, состоящим из малой компоненты

$$I_2 = 2^{1/2} A \cos 2\pi f_2 t,$$

связанной с узкой полосой частот, наложенной на большую, медленно изменяющуюся компоненту, связанную с полосой более низких частот. Так как эффективное значение второй составляющей равно $W^{1/2}$, то ей можно приписать некоторую частоту f_1 и тогда приближенно

$$I_1 = (2W)^{1/2} \cos 2\pi f_1 t.$$

Нули тока шумов сосредоточены вокруг нулей второй волны. Вблизи такого нуля

$$I_1 = \pm (2W)^{1/2} 2\pi f_1 \Delta t,$$

где Δt есть расстояние от нуля. Колебания I_1 создают нули, когда $|I_1|$ меньше, чем амплитуда I_2 , или когда

$$A > W^{1/2} 2\pi f_1 |\Delta t|,$$

а интервал, в котором возникают нули, определяется из

$$2\Delta t = \frac{AW^{-1/2}}{\pi f_1}.$$

Число нулей равно этой величине, умноженной на $2f_2$. Так как в 1 сек. $2f_1$ таких интервалов, то число нулей в 1 сек. равно

$$\frac{4}{\pi} AW^{-1/2} f_2.$$

Это выражение отличается от полученного из нашей формулы наличием множителя $\frac{2}{\pi}$. Расхождение объясняется тем, что ток шумов был представлен в двух частотных полосах в виде синусоидальных волн I_1 и I_2 .

Из этого примера ясно, что когда интеграл для ψ_0 сходится при $A \rightarrow 0$ и в то же самое время интеграл для ψ_0'' расходится при $f_2 \rightarrow \infty$ таким образом, что $Af_2 \rightarrow \infty$, ток шумов ведет себя как непрерывная функция, не имеющая производной. Кажется, для физических систем интегралы будут всегда сходиться, так как паразитные эффекты приведут к тому, что $\omega(f)$ стремится к нулю достаточно быстро. Частота, представляющая область, в которой это происходит, имеет порядок частоты микроскопических колебаний.

Ряд экспериментов показывает, что в некоторых случаях справедливость формул этого раздела нарушается. Так, если ток флуктуаций, занимающий весьма широкий диапазон, будет протекать через контур, состоящий из конденсатора C , включенного параллельно с последовательным соединением катушки индуктивности L и сопротивления R , то уравнение (3.3—11) говорит, что ожидаемое число нулей в 1 сек. тока I , протекающего через R (и L), не зависит от R . Оно просто равно $\frac{1}{\pi} (LC)^{-1/2}$. Дифференциальное уравнение для I одинаково с уравнением, которому подчиняется броуновское движение зеркальца, подвешенного в газе; давление газа играет роль R . Из кривых, описывающих это движение, видно, что их характер сильно зависит от давления. К сожалению, из кривых трудно заключить, зависит ли ожидаемое число нулей от давления. Разность между кривыми при различных давлениях указывает, что здесь должна быть некоторая зависимость¹⁾.

3.4. РАСПРЕДЕЛЕНИЕ НУЛЕЙ

Задача нахождения функции распределения для промежутков между двумя следующими друг за другом нулями представляется

¹⁾ С тех пор, как это было написано, М. Кэк и Х. Хурвиц исследовали задачу об ожидаемом числе нулей, пользуясь совсем другим методом, в котором применяется представление дробового эффекта (разд. 3.11). Их выводы подтверждают справедливость (3.3—11), когда интегралы сходятся. Если интегралы расходятся, то должно быть рассмотрено среднее число электронов в 1 сек., вызывающих дробовой эффект.

весьма трудной, и, кажется, удовлетворительного решения еще не дано. Здесь будут приведены некоторые выводы, связанные с общей проблемой, которые дают представление о форме распределения для области малых промежутков между нулями.

Будет доказано [в разделе, начиная с уравнения (3.4—12)], что вероятность прохождения тока шумов I через нуль в интервале $(\tau, \tau + d\tau)$ с отрицательной крутизной, если известно, что I проходит через нуль при $\tau=0$ с положительной крутизной, равна

$$\frac{d\tau}{2\pi} \left(\frac{\psi_0}{-\psi_0''} \right)^{1/2} (M_{22}^2 - M_{23}^2) (\psi_0^2 - \psi_\tau^2)^{-3/2} [1 + H \operatorname{arccctg}(-H)], \quad (3.4-1)$$

где M_{22} и M_{23} — алгебраические дополнения элементов $\mu_{22} = -\psi_0''$ и $\mu_{23} = -\psi_\tau''$ в матрице

$$M = \begin{vmatrix} \psi_0 & 0 & \psi_\tau' & \psi_\tau \\ 0 & -\psi_0'' & -\psi_\tau'' & -\psi_\tau' \\ \psi_\tau' & -\psi_\tau'' & -\psi_0'' & 0 \\ \psi_\tau & -\psi_\tau' & 0 & \psi_0 \end{vmatrix},$$

$$H = M_{23} (M_{22}^2 - M_{23}^2)^{-1/2}. \quad (3.4-2)$$

Выбираем $0 \leq \operatorname{arccctg}(-H) \leq \pi$, значение π берется при $\tau=0$, а $\pi/2$ при $\tau \rightarrow \infty$. Следует помнить, что аргументы функций корреляции представлены в виде индексов, т. е. $-\psi_\tau''$ в действительности равно

$$-\psi_\tau'' = -\psi''(\tau) = 4\pi^2 \int_0^\infty f^2 \omega(f) \cos 2\pi f \tau df. \quad (3.3-8)$$

По мере того как τ становится все больше и больше, на поведение I в момент τ все меньше влияет то обстоятельство, что ток проходит через нуль с положительной крутизной при $\tau=0$. Следовательно, (3.4—1) должно сходиться к вероятности того, что в некотором интервале длительностью $d\tau$, выбранном случайно, I пройдет через нуль с отрицательной крутизной. Вследствие симметрии это равно половине вероятности того, что I пройдет через нуль. Поэтому (3.4—1) должно стремиться, согласно (3.3—10), к

$$\frac{d\tau}{2\pi} \left(\frac{-\psi_0''}{\psi_0} \right)^{1/2} \quad (3.4-3)$$

при $\tau \rightarrow \infty$. Это действительно будет так, ибо M сходится к диагональной матрице, а как M_{23} , так и H стремятся к нулю при $M_{23}/H \rightarrow M_{22} \rightarrow -\psi_0''/\psi_0$. Для фильтра нижних частот с граничной частотой f_b выражение (3.4—3) дает

$$d\tau f_b 3^{-1/2} \quad (3.4-4)$$

Поведение выражения (3.4—1) при $\tau \rightarrow 0$ можно определить без больших затруднений. M_{22} и M_{23} стремятся к нулю, как τ^4 , $M_{22}^2 - M_{23}^2$, как τ^{10} , и соответственно H стремится к бесконечности как τ^{-1} . В результате (3.4—1) стремится к

$$d\tau \frac{\tau}{8} \left(\frac{\psi_0 \psi_0^{(4)} - \psi_0''^2}{-\psi_0 \psi_0''} \right) \quad (3.4-5)$$

при $\tau \rightarrow 0$, полагая, что $\psi_0^{(4)}$ существует. Здесь индекс (4) указывает на четвертую производную при $\tau=0$

$$\psi_0^{(4)} = 16\pi^4 \int_0^\infty f^4 \omega(f) df. \quad (3.4-6)$$

Для фильтра нижних частот с граничной частотой f_b выражение (3.4—5) равно

$$d\tau \frac{\tau}{30} (2\pi f_b)^2. \quad (3.4-7)$$

Если (3.4—1) применить к случаю фильтра нижних частот, то вместо τ более удобно иметь дело с переменной

$$\varphi = 2\pi f_b \tau, \quad d\varphi = 2\pi f_b d\tau. \quad (3.4-8)$$

Поэтому, если написать (3.4—1) в виде $p(\varphi) d\varphi$, то из (3.4—4) и (3.4—7) следует, что

$$p(\varphi) \rightarrow \frac{1}{2\pi\sqrt{3}} = 0,0919 \text{ при } \varphi \rightarrow \infty, \quad (3.4-9)$$

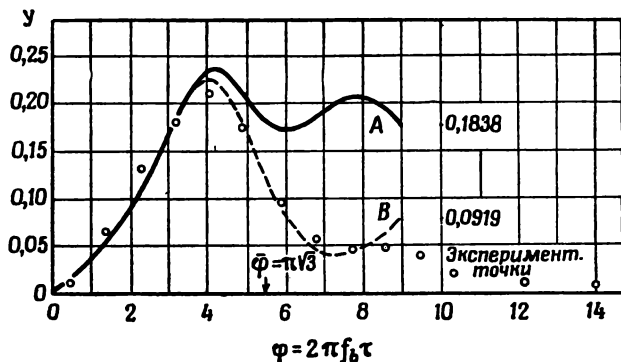
$$p(\varphi) \rightarrow \frac{\varphi}{30} \text{ при } \varphi \rightarrow 0.$$

На фиг. 1 представлена $p(\varphi)$ как функция φ для диапазона значений φ от 0 до 9. Из рассмотрения кривой и из теоретических соображений ясно, что за пределами $\varphi > 9$ функция $p(\varphi)$ колеблется вокруг 0,0919 с постоянно уменьшающейся амплитудой.

Можно принять, что $p(\varphi) d\varphi$ есть вероятность того, что I проходит через нуль в интервале $(\varphi, \varphi + d\varphi)$, если известно, что I проходит через нуль при $\varphi=0$ с крутизной, обратной по знаку крутизне при φ . $p(\varphi) d\varphi$ превышает вероятность того, что I проходит через нуль при $\varphi=0$ и в интервале $(\varphi, \varphi + d\varphi)$, не имея нулей между ними. Это объясняется тем, что $p(\varphi) d\varphi$ включает все кривые последнего класса и, кроме того, те кривые, которые могут иметь четное число нулей между 0 и φ . Отсюда следует, что кривая, представляющая плотность вероятностей интервалов между нулями, должна располагаться под кривой $p(\varphi)$.

Частичные неточности кривой $p(\varphi)$ можно обнаружить при сравнении ее с функцией плотности вероятностей, эксперименталь-

но полученной М. Е. Кэмпбеллом для интервалов между 754 следующими друг за другом нулями. Он пропускал шумы теплового движения через полосовой фильтр с нижней частотой пропускания около 200 *гц* и верхней частотой пропускания около 3000 *гц*. Ввиду плавного изгиба частотной характеристики трудно отметить точное значение верхней граничной частоты. Кружки на фиг. 1 соответ-



Фиг. 1. Распределение вероятностей интервалов между нулями тока шумов на выходе фильтра нижних частот.

$y_A \Delta\varphi$ — вероятность появления нуля в интервале $\Delta\varphi$, если в начале отсчета нуль. $y_B \Delta\varphi$ — вероятность появления нуля в интервале $\Delta\varphi$, если в начале отсчета нуль и крутизна кривых в нулевых точках противоположного знака. $y_B = p(\varphi)$, f_b — граничная частота фильтра, τ — промежуток времени между нулевыми точками.

ствуют данным Кэмпбелла, если предположить, что его фильтр ведет себя как фильтр нижних частот с частотой среза $f_b = 2850$ *гц*; последняя величина выбрана для того, чтобы максимум кривой Кэмпбелла совпадал с максимумом кривой $p(\varphi)$.

Как видно, некоторые кружки лежат выше кривой $p(\varphi)$, что, вероятно, объясняется тем, что характеристики реального фильтра в действительности отличаются от принятых нами для фильтра нижних частот.

На фиг. 1 нанесена также кривая, тесно связанная с (3.4—1). Она относится к случаю фильтра нижних частот и представляет вероятность прохождения I через нуль в интервале $(\tau, \tau + d\tau)$, когда известно, что I проходит через нуль при $\tau = 0$:

$$\frac{d\tau}{\pi} \left(\frac{\psi_0}{-\psi_0''} \right)^{1/2} \left(\frac{M_{23}}{H} \right) (\psi_0^2 - \psi_\tau^2)^{-3/2} (1 + H \arctg H), \quad (3.4-10)$$

где обозначения те же, что и в (3.4—1), а $-\frac{\pi}{2} \leq \arctg H \leq \frac{\pi}{2}$. Эта кривая расположена всегда выше кривой $p(\varphi)$, а небольшая разни-

ца между кривыми до $\varphi=4$ указывает на то, что вплоть до этой точки $\rho(\varphi)$ хорошо представляет действительное распределение нулей.

Если (3.4—1) применить к случаю полосового фильтра со сравнительно узкой полосой или к какому-нибудь подобному устройству, то можно сделать некоторые приближения и получить несколько более простое выражение, чем (3.4—1). Рассмотрим обычный идеальный полосовой фильтр с полосой пропускания от f_a до f_b . Функция корреляции и средний квадрат шумов, согласно (3.2—5), равны

$$\psi_\tau = \frac{\omega_0}{\pi\tau} \sin \pi\tau(f_b - f_a) \cos \pi\tau(f_b + f_a),$$

$$\psi_0 = \omega_0(f_b - f_a). \quad (3.2—5)$$

Из физических соображений ясно, что для случая узкополосного фильтра большинство промежутков между нулями равно приближенно

$$\tau_1 = \frac{1}{f_b + f_a},$$

т. е. примерно равно промежуткам между нулями синусоидальной волны, имеющей частоту, равную средней частоте фильтра. Поэтому можно предполагать, что (3.4—1) будет иметь пик, весьма близкий к τ_1 . Можно также ожидать пиков при $3\tau_1$, $5\tau_1$ и т. д., но они здесь не рассматриваются (исследуется только поведение (3.4—1) вблизи τ_1).

Оказывается, M_{23} примерно равно M_{22} , так что H велико, и (3.4—1) приближенно равно

$$\frac{d\tau}{2} \left(\frac{\psi_0}{-\psi_0''} \right)^{1/2} \frac{M_{23}}{(\psi_0^2 - \psi_\tau^2)^{3/2}},$$

где τ лежит вблизи τ_1 .

Чтобы показать, что M_{23} примерно равно M_{22} , следует воспользоваться выражениями

$$M_{22} = -\psi_0'' (\psi_0^2 - \psi_\tau^2) - \psi_0 \psi_\tau'^2,$$

$$M_{23} = \psi_\tau'' (\psi_0^2 - \psi_\tau^2) + \psi_\tau \psi_\tau'^2,$$

$$M_{22} + M_{23} = (\psi_0 - \psi_\tau) [(\psi_0 + \psi_\tau)(\psi_\tau'' - \psi_0'') - \psi_\tau'^2] = (\psi_0 - \psi_\tau)(B + C),$$

$$M_{22} - M_{23} = (\psi_0 + \psi_\tau) [(\psi_0 - \psi_\tau)(-\psi_\tau'' - \psi_0'') - \psi_\tau'^2] =$$

$$= (\psi_0 + \psi_\tau)(-B + C),$$

$$B = \psi_0 \psi_\tau'' - \psi_\tau \psi_0'',$$

$$C = -\psi_0 \psi_0'' + \psi_\tau \psi_\tau'' - \psi_\tau'^2$$

Из (3.2—5) видно, что ψ_τ может быть написана в виде

$$\psi_\tau = A \cos \beta\tau, \quad \beta = \pi (f_b + f_a),$$

где $\beta\tau_1 = \pi$, а A есть функция τ , медленно изменяющаяся по сравнению с $\cos \beta\tau$. Видно, что вблизи τ_1 ψ_τ примерно равна $-\psi_0$. Так же точно ψ'_τ колеблется вблизи нуля, а ψ''_τ примерно равна $-\psi''_0$. Дифференцирование по τ дает

$$\begin{aligned} \psi'_\tau &= A' \cos \beta\tau - A \beta \sin \beta\tau, \\ \psi''_\tau &= (A'' - A \beta^2) \cos \beta\tau - 2A' \beta \sin \beta\tau, \\ \psi''_0 &= A''_0 - A_0 \beta^2, \quad \psi_0 = A_0, \end{aligned}$$

где A_0 и A''_0 есть значения A и ее второй производной при $\tau=0$. Это приводит к

$$\begin{aligned} B &= (A_0 A'' - A A''_0) \cos \beta\tau - 2A_0 A' \beta \sin \beta\tau, \\ C &= (A A'' - A'^2) \cos^2 \beta\tau - A_0 A''_0 + (A_0^2 - A)^2 \beta^2. \end{aligned}$$

Покажем, что $C+B$ и $C-B$ суть величины одного порядка. Если мы это сделаем, то отсюда будет следовать, что $M_{22} - M_{23}$ значительно меньше, чем $M_{22} + M_{23}$, так как $\psi_0 - \psi_{\tau_1}$ примерно равно $2\psi_0$, а $\psi_0 + \psi_{\tau_1}$ совсем мало. Соответственно будет показано, что M_{23} почти равно M_{22} .

До сих пор не делалось приближений. Представим теперь медленно изменяющуюся функцию A в виде степенного ряда по τ . Так как для рассматриваемого типа функций ψ'_0 и ψ''_0 должны быть равны нулю, то, следовательно:

$$\begin{aligned} A &= A_0 + \frac{\tau^2}{2} A''_0 + \dots, \\ A' &= \tau A'_0 + \dots, \\ A'' &= A''_0 + \frac{\tau^2}{2} A^{(4)}_0 + \dots, \end{aligned}$$

где мы пренебрегали всеми степенями выше второй. Умножение и возведение в квадрат дают

$$\begin{aligned} A^2 - A_0^2 &= \tau^2 A_0 A''_0, \\ A A'' - A'^2 &= A_0 A''_0 + \frac{\tau^2}{2} (A_0 A^{(4)}_0 - A''_0{}^2) = A_0 A''_0 + F, \\ A_0 A'' - A A''_0 &= \frac{\tau^2}{2} (A_0 A^{(4)}_0 - A''_0{}^2) = F. \end{aligned}$$

Так как (для малых τ) A и A'' примерно равны, соответственно A_0 и A''_0 , то разность слева мала сравнительно с $A_0 A''_0$, т. е.

$$|F| \ll |A_0 A''_0|.$$

Тогда выражения для B и C приближенно равны

$$B = F \cos \beta\tau - 2A_0 A_0'' \beta\tau \sin \beta\tau,$$

$$C = F \cos^2 \beta\tau - A_0 A_0'' \sin^2 \beta\tau - A_0 A_0'' \beta^2 \tau^2.$$

Если τ находится вблизи τ_1 , то $\beta\tau$ приближенно равно π . Отсюда как $C+B$, так и $C-B$ приближенно равны $-A_0 A_0'' \pi^2$ и являются величинами одного порядка. Соответственно M_{22} и M_{23} примерно равны, а

$$M_{23} = \psi_0 (C+B) = -A_0^2 A_0'' \pi^2.$$

Если это выражение для M_{23} использовать в приближении для (3.4—1), то получим следующий результат: если функция корреляции имеет вид

$$\psi_\tau = A \cos \beta\tau,$$

где A — медленно изменяющаяся функция τ , то вероятность того, что расстояние между двумя следующими друг за другом нулями лежит между τ и $\tau+d\tau$, приближенно равна

$$\frac{d\tau}{2} \frac{a}{[1+a^2(\tau-\tau_1)^2]^{3/2}},$$

где a — положительно, и

$$a^2 = \frac{A_0 \beta^2}{-A_0'' \tau_1^2}, \quad \tau_1 = \frac{\pi}{\beta}.$$

Для идеального полосового фильтра с полосой пропускания $(f_b - f_a)$

$$a = \sqrt{3} \frac{(f_b + f_a)^2}{f_b - f_a}, \quad \tau_1 = \frac{1}{f_b + f_a},$$

а среднее значение $|\tau - \tau_1| = a^{-1}$. Поэтому

$$\frac{|\tau - \tau_1|}{\tau_1} = \frac{1}{a\tau_1} = \frac{f_b - f_a}{\sqrt{3}(f_b + f_a)} = \frac{1}{2\sqrt{3}} \frac{(\text{полоса пропускания})}{(\text{средняя частота})}.$$

Если функция корреляции не может быть представлена таким образом, но все же ведет себя подобно синусоидальной волне с медленно изменяющейся амплитудой, то можно использовать первое приближение для выражения (3.4—1). Тогда, вероятность того, что расстояние между двумя следующими друг за другом нулями лежит между τ и $\tau+d\tau$, приближенно равна

$$\frac{bd\tau}{(\psi_0^2 - \psi_\tau^2)^{3/2}},$$

когда τ лежит вблизи τ_1 , где τ_1 есть то наименьшее значение τ , при котором $\psi_\tau \approx -\psi_0$. Эта вероятность предполагается быстро сходящейся к нулю при τ , удаляющемся от τ_1 , а b выбирается так, что интеграл по действующей области вокруг τ_1 равен единице.

Наиболее трудной задачей является, повидимому, получение выражения для распределения нулей в случае больших расстояний между нулями. Первый метод заключается в расширении условий, приводящих к уравнению (3.4—1), путем добавления условия, чтобы I было положительным в равноотстоящих точках вдоль оси времени между 0 и τ . Этот метод приводит к трудно вычислимым интегралам; для одной точки между 0 и τ интеграл имеет вид (3.5—7).

Другой метод исследования состоит в применении способа «включения и исключения» нулей между 0 и τ .

Рассмотрим класс кривых I , имеющих нуль при $\tau=0$. Тогда теоретически можно будет найти функции $p_0(\tau)$, $p_1(r, \tau)$, $p_2(r, s, \tau)$, связанные с этим классом, где $p_0(\tau) d\tau$ — вероятность того, что кривая имеет нуль в интервале $d\tau$.

$p_1(r, \tau) dr d\tau$ — вероятность того, что кривая имеет нули в интервалах $d\tau$ и dr .

$p_2(r, s, \tau) dr ds d\tau$ — вероятность того, что кривая имеет нули в интервалах $d\tau$, ds и dr .

Выражение для $p_0(\tau) d\tau$ дается уравнением (3.4—10). Метод «включения и исключения» приводит тогда к выражению для $P_0(\tau) d\tau$ — вероятности получения нуля при 0 и нуля в интервале $(\tau, \tau+d\tau)$, но отсутствия нулей между 0 и τ . Эта вероятность равна

$$P_0(\tau) = p_0(\tau) - \frac{1}{1!} \int_0^\tau p_1(r, \tau) dr + \frac{1}{2!} \int_0^\tau \int_0^\tau p_2(r, s, \tau) dr ds - \\ - \frac{1}{3!} \int_0^\tau \int_0^\tau \int_0^\tau p_3(r, s, t, \tau) dr ds dt + \dots \quad (3.4-11)$$

Выражение (3.4—11) может быть применено для случайных событий, происходящих независимо. Поэтому, если $\nu d\tau$ есть вероятность наступления события в интервале $d\tau$, ν постоянно, а события независимы, то получим для p_0, p_1, p_2, \dots соответственно ν, ν^2, ν^3, \dots . Из (3.4—11) приходим к известному выводу, что $P_0(\tau) = \nu e^{-\nu\tau}$.

Докажем теперь (3.4—1). Вывод основан на обобщении (3.3—5): если y есть случайная кривая, описываемая (3.3—1), то вероятность того, что y пройдет через нуль в интервале $(x_1, x_1 + dx_1)$ с положительной крутизной и через нуль в интервале $(x_2, x_2 + dx_2)$ с отрицательной крутизной, равна

$$- dx_1 dx_2 \int_0^\infty d\eta_1 \int_{-\infty}^0 d\eta_2 \eta_1 \eta_2 p(0, \eta_1, x_1; 0, \eta_2, x_2), \quad (3.4-12)$$

где $p(\xi_1, \eta_1, x_1; \xi_2, \eta_2, x_2)$ есть функция плотности вероятностей для четырех случайных переменных

$$\xi_i = F(a_1, a_2, \dots, a_N; x_i),$$

$$\eta_i = \left(\frac{\partial F}{\partial x} \right)_{x=x_i}, \quad i=1,2.$$

В уравнении (3.4—12) x_1 и x_2 играют роль параметров. Этот вывод может быть найден во многом аналогично способу получения (3.3—5).

Если установить идентичность F с одним из представлений тока шумов $I(t)$, т. е. с (2.8—1), либо с (2.8—6), то видно, что p подчиняется нормальному закону во всех четырех измерениях. Можно получить моменты второго порядка непосредственно из этого представления, как это было сделано в уравнениях, приведенных вслед за (3.3—7). Тот же самый результат можно получить из определения $\psi(\tau)$ и, для разнообразия, можно выбрать этот второй метод. Положим $x_1 = t_1$, $x_2 = t_1 + \tau$. Тогда

$$\overline{\xi_1^2} = \overline{\xi_2^2} = \overline{I^2(t)} = \psi_0,$$

$$\overline{\xi_1 \xi_2} = \overline{I(t) I(t + \tau)} = \psi_\tau, \quad (3.4—13)$$

$$\overline{\eta_1 \eta_2} = \left(\frac{\partial I}{\partial t} \right)_t \left(\frac{\partial I}{\partial t} \right)_{t+\tau} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T I'(t + \tau) I'(t) dt,$$

где штрихи обозначают дифференцирование по аргументам. Интегрируя по частям, найдем

$$\int_0^T I'(t + \tau) dI(t) = [I'(t + \tau) I(t)]_0^T - \int_0^T I''(t + \tau) I(t) dt.$$

Полагаем, что I и ее производные остаются конечными, так что при делении на T интегрируемые части в пределе исчезают. Так как

$$I''(t + \tau) = \frac{\partial^2}{\partial \tau^2} I(t + \tau),$$

то имеем

$$\overline{\eta_1 \eta_2} = - \frac{\partial^2}{\partial \tau^2} \psi(\tau) = - \psi''_\tau.$$

Полагая $\tau=0$, найдем

$$\overline{\eta_1^2} = \overline{\eta_2^2} = - \psi''_0$$

в согласии со значением μ_{22} , полученным из (3.3—7). Таким же образом

$$\overline{\xi_1 \eta_2} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T I'(t + \tau) I(t) dt = \frac{\partial}{\partial \tau} \psi(\tau) = \psi'_\tau,$$

$$\overline{\xi_2 \eta_1} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T I'(t) I(t + \tau) dt = \lim_{T \rightarrow \infty} (-) \frac{1}{T} \int_0^T I'(t + \tau) I(t) dt = - \psi'_\tau,$$

где для получения $\overline{\xi_2 \eta_1}$ произведено интегрирование по частям. Полагая $\tau=0$ и используя $\psi'_0=0$, найдем

$$\overline{\xi_1 \eta_1} = \overline{\xi_2 \eta_2} = 0.$$

Чтобы получить матрицу M моментов второго порядка μ_{rs} в виде совершенно симметричном относительно ее центра, выбираем наши переменные 1, 2, 3, 4 в порядке $\xi_1, \eta_1, \eta_2, \xi_2$. Из уравнений (3.4—13) и последующего видим, что этот выбор приводит к выражению (3.4—2) для M .

Если положить ξ_1 и ξ_2 равными нулю, то получим для функции плотности вероятностей в (3.4—12) выражение

$$\frac{|M|^{-1/2}}{4\pi^2} \exp \left[-\frac{1}{2|M|} (M_{22}\eta_1^2 + 2M_{23}\eta_1\eta_2 + M_{33}\eta_2^2) \right].$$

Вследствие симметрии матрицы M , $M_{22}=M_{33}$. Если в интеграле (3.4—12) произвести замену переменных

$$x = \left[\frac{M_{22}}{2|M|} \right]^{1/2} \eta_1, \quad y = - \left[\frac{M_{22}}{2|M|} \right]^{1/2} \eta_2,$$

то получим

$$\frac{dx_1 dx_2}{\pi^2} \frac{|M|^{3/2}}{M_{22}^2} \int_0^\infty x dx \int_0^\infty dy y e^{-x^2 - y^2 + 2(M_{23}/M_{22})xy}.$$

Двойной интеграл можно вычислить при помощи (3.5—4). Пусть

$$\varphi = \arccos \left(-\frac{M_{23}}{M_{22}} \right) = \operatorname{arctg}(-H), \quad H = M_{23} (M_{22}^2 - M_{23}^2)^{-1/2},$$

где H имеет то же значение, что и в (3.4—2). Тогда выражение превращается в

$$\frac{dx_1 dx_2}{4\pi^2} \frac{|M|^{3/2}}{M_{22}^2 - M_{23}^2} [1 + H \operatorname{arctg}(-H)].$$

По свойству определителей

$$M_{22}M_{33} - M_{23}^2 = |M| (\psi_0^2 - \psi_\tau^2).$$

Используем это соотношение для исключения $|M|$ и разделим выражение на

$$\frac{dx_1}{2\pi} \left(\frac{-\psi_0''}{\psi_0} \right)^{1/2}$$

что, как это следует из (3.3—10), есть вероятность прохождения через нуль в интервале $(x_1, x_1 + dx_1)$ с положительной крутизной.

В результате получим вероятность прохождения через нуль в интервале dx_2 с отрицательной крутизной, когда известно, что I

проходит через нуль при x_{τ} с положительной крутизной:

$$\frac{dx_2}{2\pi} \left(\frac{\psi_0}{-\psi_0''} \right)^{1/2} (M_{22}^2 - M_{23}^2)^{1/2} (\psi_0^2 - \psi_{\tau}^2)^{-3/2} [1 + H \operatorname{arccctg}(-H)].$$

Это и есть (3.4—1).

Выражение (3.4—10) дает также вероятность прохождения I через нуль в интервале $d\tau$, когда известно, что I проходит через нуль вначале с положительной крутизной. Эта вероятность может быть получена из (3.4—1) путем добавления вероятности прохождения I через нуль в интервале $d\tau$ с положительной крутизной, если известно, что этот ток проходит через нуль с положительной крутизной. Поэтому нужно добавить выражение, содержащее интеграл, в котором интегрирование по отношению к η_1 и η_2 производится в пределах от 0 до ∞ . Этот интеграл, написанный при помощи введения переменных x и y , равен

$$\int_0^{\infty} x dx \int_0^{\infty} dy y e^{-x^2 - y^2 - 2(M_{23}/M_{22})xy}.$$

Это эквивалентно изменению знака M_{23} и, следовательно, и H . После сложения надо рассмотреть

$$1 + H \operatorname{arccctg}(-H) + 1 - H \operatorname{arccctg} H = 2 + H [\operatorname{arccctg}(-H) - \operatorname{arccctg} H] = 2 + H (\pi - 2 \operatorname{arccctg} H) = 2 (1 + H \operatorname{arccctg} H),$$

а это приводит к (3.4—10).

3.5. КРАТНЫЕ ИНТЕГРАЛЫ, ВХОДЯЩИЕ В ФОРМУЛЫ

Нужно вычислить интегралы вида

$$J = \int_0^{\infty} dx_1 \int_0^{\infty} dx_2 e^{-x_1^2 - 2ax_1x_2 - x_2^2}. \quad (3.5—1)$$

Одним из методов является сведение показательной функции к сумме квадратов путем соответствующей линейной замены переменной, а затем преобразование к полярным координатам. Этот метод пригоден также для тройных интегралов такого же типа, но если применить его к четырехкратным интегралам, то последнее интегрирование, повидимому, не может быть выполнено в конечной форме.

Сведение показательной функции к сумме квадратов основано на следующем преобразовании. Если

$$\begin{aligned} x_1 &= h_1 y_1 + h_2 D_{21} y_2 + h_3 D_{31} y_3 + \dots + h_n D_{n,1} y_n, \\ x_2 &= 0 + h_2 D_{22} y_2 + \dots + h_n D_{n,2} y_n, \\ &\vdots \\ x_n &= 0 + 0 + \dots + h_n D_{n,n} y_n, \end{aligned} \quad (3.5—2)$$

где $D_0=1$, $D_1=a_{11}$, $D_{r,r}=D_{r-1}$, а D_{rs} есть алгебраическое дополнение элемента a_{sr} (или a_{rs} , так как они равны) в D_r :

$$D_r = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1r} \\ a_{12} & a_{22} & & \\ \dots & & & \\ a_{1r} & \dots & & a_{rr} \end{vmatrix}, \quad h_r = (D_{r-1} D_r)^{-1/2},$$

тогда, если ни один из D_r не равен нулю:

$$\sum_1^n a_{rs} x_r x_s = y_1^2 + y_2^2 + \dots + y_n^2.$$

Из (3.5—2) якобиан $\frac{\partial(x_1, \dots, x_n)}{\partial(y_1, \dots, y_n)}$ равен $D_n^{-1/2}$.

Применяем это преобразование к показательной функции

$$x_1 = y_1 - a D_2^{-1/2} y_2,$$

$$x_2 = 0 + D_2^{-1/2} y_2,$$

$$D_2 = 1 - a^2.$$

Так как x_2 пробегает от 0 до ∞ , то так же должен вести себя и y_2 . Выражение для x_1 показывает, что y_1 пробегает от $a D_2^{-1/2} y_2$ до ∞ . Поэтому интеграл равен

$$J = D_2^{-1/2} \int_0^\infty dy_2 \int_{a D_2^{-1/2} y_2}^\infty e^{-y_1^2 - y_2^2} dy_1.$$

Перейдем теперь к полярным координатам

$$y_1 = \rho \cos \theta, \quad y_2 = \rho \sin \theta,$$

$$dy_1 dy_2 = \rho d\rho d\theta,$$

$$y_2 \geq 0 \text{ при } 0 \leq \theta \leq \pi,$$

$$y_1 \geq a D_2^{-1/2} y_2 \text{ при } \operatorname{ctg} \theta \geq a D_2^{-1/2}$$

и получим

$$J = D_2^{-1/2} \int_0^{\operatorname{arccctg}(a D_2^{-1/2})} d\theta \int_0^\infty \rho e^{-\rho^2} d\rho = \frac{1}{2} D_2^{-1/2} \operatorname{arccctg}(a D_2^{-1/2}),$$

где $\operatorname{arccctg}$ лежит между 0 и π . Можно написать и в более простой форме

$$J = \frac{1}{2} (1 - a^2)^{-1/2} \arccos a = \frac{1}{2} \varphi \operatorname{cosec} \varphi,$$

где

$$a = \cos \varphi,$$

причем подразумевается, что $0 \leq \varphi \leq \pi$.

Другие интегралы можно получить путем дифференцирования. Поэтому из

$$\int_0^{\infty} dx \int_0^{\infty} dy e^{-x^2-y^2-2xy \cos \varphi} = \frac{1}{2} \varphi \operatorname{cosec} \varphi \quad (3.5-3)$$

получим

$$\int_0^{\infty} dx \int_0^{\infty} dy xy e^{-x^2-y^2-2xy \cos \varphi} = \frac{1}{4} \operatorname{cosec}^2 \varphi (1 - \varphi \operatorname{ctg} \varphi). \quad (3.5-4)$$

Используя такое же преобразование, можем найти

$$\int_0^{\infty} dx \int_0^{\infty} dy ye^{-x^2-y^2-2axy} = \frac{\sqrt{\pi}}{4} \frac{1}{1+a}. \quad (3.5-5)$$

Конечно, можно разложить показательную часть в степенной ряд и интегрировать почленно, но это приводит к ряду, который должен быть суммирован в каждом частном случае:

$$\int_0^{\infty} dx \int_0^{\infty} dy x^n y^m e^{-x^2-y^2-2axy} = \frac{1}{4} \sum_{r=0}^{\infty} \frac{(-2a)^r}{r!} \Gamma\left(\frac{n+r+1}{2}\right) \Gamma\left(\frac{m+r+1}{2}\right).$$

Если принять $-1 < \operatorname{Re}(n) < -\frac{1}{2}$, $-1 < \operatorname{Re}(m) < -\frac{1}{2}$, то ряды могут быть суммированы при $a=1$. Результат, сформулированный в соответствующем месте текста после уравнения (3.8—9), получен путем аналитического продолжения m и n .

Тот же метод применим, если пределы будут $\pm\infty$. Когда m и n целые числа, то получим

$$\int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy x^n y^m e^{-x^2-y^2-2xy \cos \varphi} = \begin{cases} 0, & n+m - \text{нечетное} \\ (-1)^n \sqrt{\pi} \frac{\Gamma\left(\frac{m+n+1}{2}\right)}{(\sin \varphi)^{n+m+1}} F\left(-n, -m; \frac{1-n-m}{2}; \frac{1-\cos \varphi}{2}\right), & n+m - \text{четное} \end{cases} \quad (3.5-6)$$

Гипергеометрическая функция может быть также написана в виде

$$F\left(-\frac{n}{2}, -\frac{m}{2}; \frac{1-n-m}{2}; \sin^2 \varphi\right).$$

Путем преобразования этого выражения приходим к следующим значениям для интеграла:

0, $n+m$ — нечетное,

$$\frac{\Gamma\left(\frac{m+1}{2}\right) \Gamma\left(\frac{n+1}{2}\right)}{(\sin \varphi)^{n+m+1}} F\left(-\frac{n}{2}, -\frac{m}{2}, \frac{1}{2}; \cos^2 \varphi\right), \quad m, n \text{ — четные,}$$

$$-2 \frac{\Gamma\left(1 + \frac{n}{2}\right) \Gamma\left(1 + \frac{m}{2}\right)}{(\sin \varphi)^{n+m+1}} \cos \varphi F\left(\frac{1-m}{2}, \frac{1-n}{2}; \frac{3}{2}; \cos^2 \varphi\right),$$

m, n — нечетные.

Как указывалось раньше, метод, использованный для вычисления двойных интегралов, может быть также применен в случае схожих тройных интегралов. Здесь приведем два результата, полученные таким образом:

$$\int_0^\infty dx \int_0^\infty dy \int_0^\infty dz \exp(-x^2 - y^2 - z^2 - 2cxy - 2bzx - 2ayz) = \\ = \frac{1}{4} \left(\frac{\pi}{D_3} \right)^{1/2} (\alpha + \beta + \gamma - \pi),$$

$$\int_0^\infty dx \int_0^\infty dy \int_0^\infty dz yz \exp(-x^2 - y^2 - z^2 - 2cxy - 2bzx - 2ayz) = \\ = \frac{\sqrt{\pi}}{8D_3} \left[\frac{1+a-b-c}{1+a} - \frac{a-bc}{D_3^{1/2}} (\alpha + \beta + \gamma - \pi) \right], \quad (3.5-7)$$

где β и γ получаются путем циклической перестановки a, b, c из

$$\alpha = \arccos \frac{a-bc}{(1-c^2)^{1/2}(1-b^2)^{1/2}} = \arcsin \left[\frac{D_3}{(1-c^2)(1-b^2)} \right]^{1/2} = \\ = \operatorname{arccctg} \frac{a-bc}{D_3^{1/2}},$$

где α, β, γ все лежат в интервале $(0, \pi)$ и где

$$D_3 = \begin{vmatrix} 1 & c & b \\ c & 1 & a \\ b & a & 1 \end{vmatrix} = 1 + 2abc - a^2 - b^2 - c^2.$$

Для справок приведем интегралы, которые получаются из определения нормального распределения, данного в разделе (2.9):

$$\int_{-\infty}^{+\infty} dx_1 \dots \int_{-\infty}^{+\infty} dx_n \exp\left(-\sum_1^n a_{rs} x_r x_s\right) = \left(\frac{\pi^n}{|a|}\right)^{1/2}, \\ \int_{-\infty}^{+\infty} dx_1 \dots \int_{-\infty}^{+\infty} dx_n x_r x_u \exp\left(-\sum_1^n a_{rs} x_r x_s\right) = \left(\frac{\pi^n}{|a|^3}\right)^{1/2} \frac{A_{tu}}{2}, \quad (3.5-8)$$

где квадратичная форма является определенно положительной, $|a|$ — ее определитель, A_{tu} — алгебраическое дополнение элемента a_{tu} . Кстати, эти интегралы могут рассматриваться как частные случаи следующего выражения:

$$\begin{aligned} & \int_{-\infty}^{+\infty} dx_1 \dots \int_{-\infty}^{+\infty} dx_n f\left(\sum_1^n a_{rs} x_r x_s\right) F\left(\sum_1^n b_r x_r\right) = \\ & = \frac{2}{\Gamma\left(\frac{n-1}{2}\right)} \left[\frac{\pi^{n-1}}{|a|}\right]^{1/2} \int_{-\infty}^{+\infty} dx \int_0^{\infty} dy y^{n-2} f(x^2+y^2) \times \\ & \quad \times F\left\{x \left[\frac{\sum_1^n A_{rs} b_r b_s}{|a|}\right]^{1/2}\right\}. \end{aligned} \quad (3.5-9)$$

3.6. РАСПРЕДЕЛЕНИЕ ВЫБРОСОВ ТОКА ШУМОВ

Здесь воспользуемся выводом, подобным применявшемуся в разделах 3.3 и 3.4. Пусть y будет случайная кривая, заданная уравнением (3.3—1):

$$y = F(a_1, \dots, a_N; x). \quad (3.3-1)$$

Если соответствующие условия удовлетворены, то вероятность того, что y имеет максимум в прямоугольнике $(x_1, x_1+dx_1, y_1, y_1+dy_1)$, причем dx_1 и dy_1 — величины одного порядка, равна

$$- dx_1 dy_1 \int_{-\infty}^0 p(y_1, 0, \zeta) \zeta d\zeta, \quad (3.6-1)$$

а ожидаемое число максимумов y в интервале $a \leq x \leq b$ получается путем интегрирования этого выражения по интервалу $-\infty \leq y_1 \leq \infty$, $a \leq x_1 \leq b$. Здесь $p(\xi, \eta, \zeta)$ есть функция плотности вероятностей для случайных переменных

$$\begin{aligned} \xi &= F(a_1, \dots, a_N; x_1), \\ \eta &= \left(\frac{\partial F}{\partial x}\right)_{x=x_1} \\ \zeta &= \left(\frac{\partial^2 F}{\partial x^2}\right)_{x=x_1} \end{aligned} \quad (3.6-2)$$

Применяя этот вывод как и раньше, заменим x и y на t и I . Тогда

$$\begin{aligned} \xi &= I = \sum_1^N c_n \cos(\omega_n t - \varphi_n), \\ \eta &= I', \\ \zeta &= I'', \end{aligned}$$

где штрихи обозначают дифференцирование по t . Согласно центральной предельной теореме, распределение ξ, η, ζ сходится к нормальному закону. Моменты второго порядка, определяющие этот закон, могут быть получены либо из приведенных выше определений ξ, η, ζ , либо из функции корреляции, как это было сделано вслед за уравнением (3.4—13):

$$\begin{aligned}\bar{\xi}^2 &= \psi_0, & \bar{\eta}^2 &= -\psi_0'', & \bar{\xi}\eta &= 0, \\ \bar{\eta}\zeta &= \overline{I'(t)I''(t)} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T I'(t) I''(t) dt = \\ &= \lim_{T \rightarrow \infty} \frac{1}{2T} \left[I'^2(T) - I'^2(0) \right] = 0, \\ \bar{\xi}\zeta &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T I(t) I''(t) dt = \lim_{\tau \rightarrow 0} \frac{\partial^2 \psi(\tau)}{\partial \tau^2} = \psi_0'', \\ \bar{\zeta}^2 &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T I''(t) I''(t) dt = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T I^{(4)}(t) I(t) dt = \psi_0^{(4)},\end{aligned}$$

где значок (4) обозначает четвертую производную. Матрица моментов поэтому равна

$$M = \begin{vmatrix} \psi_0 & 0 & \psi_0'' \\ 0 & -\psi_0'' & 0 \\ \psi_0'' & 0 & \psi_0^{(4)} \end{vmatrix}$$

Определитель $|M|$ и алгебраические дополнения равны

$$\begin{aligned}|M| &= -\psi_0'' (\psi_0 \psi_0^{(4)} - \psi_0''^2), \\ M_{11} &= -\psi_0'' \psi_0^{(4)}, & M_{13} &= \psi_0''^2, & M_{33} &= -\psi_0'' \psi_0.\end{aligned}\quad (3.6-3)$$

Функция плотности вероятностей в (3.6—1) равна

$$\begin{aligned}p(I, 0, \zeta) &= (2\pi)^{-3/2} |M|^{-1/2} \exp \left[-\frac{1}{2|M|} \left(M_{11} I^2 + \right. \right. \\ &\quad \left. \left. + M_{33} \zeta^2 + 2M_{13} I\zeta \right) \right]\end{aligned}\quad (3.6-4)$$

и если подставить ее в (3.6—1) и выполнить интегрирование по ζ , то получим

$$\begin{aligned}dI dt \frac{(2\pi)^{-3/2}}{M_{33}} \left[|M|^{1/2} e^{-M_{11} I^2/2|M|} + \right. \\ \left. + M_{13} I \left(\frac{\pi}{2M_{33}} \right)^{1/2} e^{-I^2/2\psi_0} \left(1 + \operatorname{erf} \frac{M_{13} I}{(2|M| M_{33})^{1/2}} \right) \right]\end{aligned}\quad (3.6-5)$$

где erf означает функцию ошибок. Это и есть вероятность наступления выброса в прямоугольнике $dl dt$.

Как указывалось в соответствующем месте текста после уравнения (3.6—1), ожидаемое число выбросов в интервале (t_1, t_2) может быть найдено путем интегрирования (3.6—1) в пределах от t_1 до t_2 после замены x на t и в пределах от $-\infty$ до $+\infty$ после замены y на I . Если применить (3.6—4), то легче сначала выполнить интегрирование по I . Тогда получим

$$\begin{aligned} & - \int_{t_1}^{t_2} dt \frac{M_{11}^{-1/2}}{2\pi} \int_{-\infty}^0 \zeta \exp \left[-\frac{\zeta^2}{2|M|} \left(M_{33} - \frac{M_{13}^2}{M_{11}} \right) \right] d\zeta = \\ & = (t_2 - t_1) \frac{\psi_0^{(4)}}{2\pi} M_{11}^{-1/2} = \frac{t_2 - t_1}{2\pi} \left[\frac{\psi_0^{(4)}}{-\psi_0''} \right]^{1/2} \end{aligned}$$

Отсюда ожидаемое число выбросов в 1 сек.

$$\frac{1}{2\pi} \left[\frac{\psi_0^{(4)}}{-\psi_0''} \right]^{1/2} = \left[\frac{\int_0^\infty f^4 \omega(f) df}{\int_0^\infty f^2 \omega(f) df} \right]^{1/2} \quad (3.6-6)$$

Для полосового фильтра ожидаемое число выбросов в 1 сек. равно

$$\left[\frac{3}{5} \frac{f_b^5 - f_a^5}{f_b^3 - f_a^3} \right]^{1/2} \quad (3.6-7)$$

где f_b и f_a — частоты среза. Полагая $f_a = 0$, найдем число выбросов для фильтра нижних частот

$$f_b \left(\frac{3}{5} \right)^{1/2} = 0,775 f_b. \quad (3.6-8)$$

Из (3.6—8) и (3.6—5) можно получить функцию плотности вероятностей для выбросов в случае фильтра нижних частот. Поэтому вероятность того, что случайно выбранный из всех выбросов будет лежать в промежутке $I, I + dI$, равна

$$\frac{dI}{3\sqrt{2\pi\psi_0}} \left\{ 2e^{-9y^2/8} + \left(\frac{5\pi}{2} \right)^{1/2} ye^{-y^2/2} \left[1 + \operatorname{erf} y \left(\frac{5}{8} \right)^{1/2} \right] \right\} \quad (3.6-9)$$

где

$$y = \frac{I}{\sqrt{\psi_0}}.$$

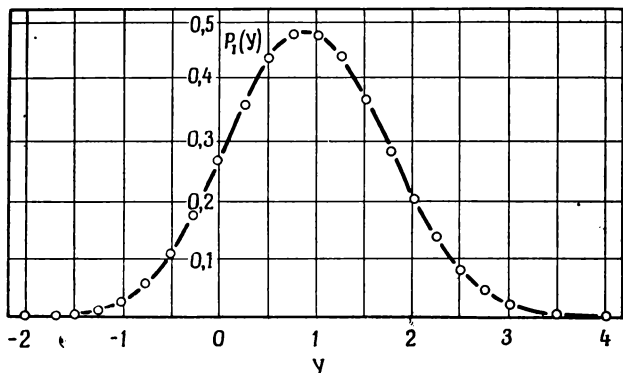
Когда y велико и положительно, (3.6—9) приближенно равно

$$\frac{dI}{\sqrt{\psi_0}} \frac{\sqrt{5}}{3} ye^{-y^2/2}$$

Если написать (3.6—9) в виде $p_I(y) dy$, то плотность вероятностей $p_I(y)$ для y может быть графически представлена фиг. 2. Функция распределения $P(I_{\text{макс.}} < y \sqrt{\psi_0})$ определяется как

$$P(I_{\text{макс.}} < y \sqrt{\psi_0}) = \int_{-\infty}^y p_I(y) dy;$$

она дает вероятность того, что случайно выбранный выброс меньше, чем заданная величина $y \sqrt{\psi_0} = I$. Это одна из четырех кривых, представленных на фиг. 4.



Фиг. 2. Плотность вероятностей выбросов тока шумов на выходе идеального фильтра нижних частот.

$$\frac{p_I(y)}{\sqrt{\psi_0}} dI - \text{вероятность} \quad \text{выбранный произвольно} \\ I_{\text{макс.}} \text{ заключен между } I \text{ и } I+dI.$$

Если I велико и положительно, то можно получить приближение из уравнения (3.6—5). Замечаем, что

$$\frac{M_{11}}{|M|} = \frac{\psi_0^{(4)}}{\psi_0 \psi_0^{(4)} - \psi_0''^2} > \frac{1}{\psi_0}$$

так что когда I велико и положительно, то

$$e^{-M_{11} I^2/2|M|} \ll e^{-I^2/2\psi_0}.$$

В этих условиях $1 + \text{erf}$ примерно равно двум. Поэтому, сохраняя только важнейшие члены и пользуясь определениями M , найдем приближение для (3.6—5)

$$\frac{dI dt}{2\pi\psi_0} \left(\frac{-\psi_0''}{\psi_0} \right)^{1/2} I e^{-I^2/2\psi_0} \quad (3.6-10)$$

Отсюда следует, что ожидаемое число выбросов в 1 сек., лежа-

щих выше линии $I=I_1$, приближенно равно, когда I_1 велико:

$$\frac{1}{2\pi} \left(\frac{-\psi_0''}{\psi_0} \right)^{1/2} e^{-I_1^2/2\psi_0} = \quad (3.6-11)$$

$$= e^{-I_1^2/2\psi_0} \frac{1}{2} \quad (\text{ожидаемое число нулей } I \text{ в } 1 \text{ сек.}).$$

Интересно отметить, что приближение (3.6—11) для ожидаемого числа выбросов, превышающих I_1 , совпадает с точным выражением (3.3—14) для ожидаемого числа моментов, в которые I проходит через I_1 с положительной крутизной.

3.7. ПЛОТНОСТЬ ВЕРОЯТНОСТЕЙ ДЛЯ ОГИБАЮЩЕЙ ТОКА ШУМОВ

Ток шумов на выходе сравнительно узкополосного фильтра имеет характер синусоидальной волны, частота которой грубо равна средней частоте полосы пропускания, а амплитуда подвержена нерегулярным флуктуациям, причем быстрота флуктуаций имеет порядок ширины полосы пропускания. Здесь будут рассмотрены флуктуации огибающей подобной волны.

Прежде всего дадим определение огибающей. Пусть f_m представляет среднюю частоту полосы пропускания фильтра. Тогда если

$$\omega_m = 2\pi f_m, \quad (3.7-1)$$

то ток шумов может быть представлен как (см. 2.8—6)

$$I = \sum_{n=1}^N c_n \cos(\omega_n t - \omega_m t - \varphi_n + \omega_m t) = \quad (3.7-2)$$

$$= I_c \cos \omega_m t - I_s \sin \omega_m t,$$

где составляющие I_c и I_s суть

$$I_c = \sum_{n=1}^N c_n \cos(\omega_n t - \omega_m t - \varphi_n),$$

$$I_s = \sum_{n=1}^N c_n \sin(\omega_n t - \omega_m t - \varphi_n). \quad (3.7-3)$$

Огибающая R есть функция t

$$R = (I_c^2 + I_s^2)^{1/2}. \quad (3.7-4)$$

Как это следует из центральной предельной теоремы и из определений (3.7—3) для I_c и I_s , здесь имеются две нормально распределенные случайные переменные. Они независимы, так как

$\overline{I_c I_s} = 0$, и имеют одинаковое стандартное отклонение, равное корню квадратному из

$$\overline{I_c^2} = \overline{I_s^2} = \overline{I^2} = \int_0^{\infty} \omega(f) df = \psi_0. \quad (3.7-5)$$

Соответственно вероятность того, что точка (I_c, I_s) лежит внутри элементарного прямоугольника $(dI_c dI_s)$, равна

$$\frac{dI_c dI_s}{2\pi\psi_0} \exp\left(-\frac{I_c^2 + I_s^2}{2\psi_0}\right) \quad (3.7-6)$$

В дальнейшем удобно ввести другую случайную переменную θ , где

$$I_c = R \cos \theta, \quad I_s = R \sin \theta. \quad (3.7-7)$$

Так как I_c и I_s — случайные переменные, то таковы же R и θ . Дифференциалы связаны так:

$$dI_c dI_s = R d\theta dR, \quad (3.7-8)$$

а функция распределения для R и θ получается из (3.7-6), если произвести замену переменных:

$$\frac{d\theta}{2\pi} \frac{R dR}{\psi_0} e^{-R^2/2\psi_0} \quad (3.7-9)$$

Так как эта функция может быть представлена как произведение членов, содержащих только R и только θ , то R и θ — независимые случайные переменные; θ равномерно распределена в интервале $(0 \div 2\pi)$, а R имеет плотность вероятностей

$$\frac{R}{\psi_0} e^{-R^2/2\psi_0} \quad (3.7-10)$$

Выражение (3.7-10) представляет собой плотность вероятностей для огибающей. Подобно нормальному закону для мгновенных значений I , она зависит только от средней полной мощности

$$\psi_0 = \int_0^{\infty} \omega(f) df.$$

Рассмотрим теперь корреляцию между R в момент t и ее значением в некоторый более поздний момент $t + \tau$. Пусть индексы 1 и 2 относятся, соответственно, к моментам t и $t + \tau$. Тогда из (3.7-3) и из центральной предельной теоремы следует, что четыре случайные переменные $I_{c_1}, I_{s_1}, I_{c_2}, I_{s_2}$ имеют нормальное распределение в четырех измерениях. Это распределение определяется моментами второго порядка

$$\begin{aligned} \overline{I_{c_1}^2} &= \overline{I_{s_1}^2} = \overline{I_{c_2}^2} = \overline{I_{s_2}^2} = \psi_0 = \mu_{11}, \\ \overline{I_{c_1} I_{s_1}} &= \overline{I_{c_2} I_{s_2}} = 0, \\ \overline{I_{c_1} I_{c_2}} &= \overline{I_{s_1} I_{s_2}} = \frac{1}{2} \sum_{n=1}^N c_n^2 \cos(\omega_n \tau - \omega_m \tau) \rightarrow \\ &\rightarrow \int_0^{\infty} \omega(f) \cos 2\pi(f - f_m)\tau df = \mu_{13}, \\ \overline{I_{c_1} I_{s_2}} &= -\overline{I_{c_2} I_{s_1}} = \frac{1}{2} \sum_{n=1}^N c_n^2 \sin(\omega_n \tau - \omega_m \tau) \rightarrow \\ &\rightarrow \int_0^{\infty} \omega(f) \sin 2\pi(f - f_m)\tau df = \mu_{14}. \end{aligned} \quad (3.7-11)$$

Матрица моментов для переменных, расположенных в порядке $I_{c_1}, I_{s_1}, I_{c_2}, I_{s_2}$, равна

$$M = \begin{vmatrix} \psi_0 & 0 & \mu_{13} & \mu_{14} \\ 0 & \psi_0 & \mu_{14} & \mu_{13} \\ \mu_{13} & \mu_{14} & \psi_0 & 0 \\ \mu_{14} & \mu_{13} & 0 & \psi_0 \end{vmatrix},$$

отсюда алгебраические дополнения определителя $|M|$ равны

$$\begin{aligned} M_{11} &= M_{22} = M_{33} = M_{44} = \psi_0 (\psi_0^2 - \mu_{13}^2 - \mu_{14}^2) = \psi_0 A, \\ A &= \psi_0^2 - \mu_{13}^2 - \mu_{14}^2, \\ M_{12} &= M_{34} = 0, \\ M_{13} &= M_{24} = -\mu_{13} A, \\ M_{14} &= -M_{23} = -\mu_{14} A, \\ |M| &= A^2. \end{aligned} \quad (3.7-12)$$

Поэтому плотность вероятностей для четырех случайных переменных равна:

$$\frac{1}{4\pi^2 A} \exp \left\{ -\frac{1}{2A} \left[\psi_0 (I_1^2 + I_2^2 + I_3^2 + I_4^2) - \right. \right. \\ \left. \left. - 2\mu_{13} (I_1 I_3 + I_2 I_4) - 2\mu_{14} (I_1 I_4 - I_2 I_3) \right] \right\},$$

где написано I_1, I_2, I_3, I_4 вместо $I_{c_1}, I_{s_1}, I_{c_2}, I_{s_2}$

Сделаем теперь преобразование

$$\begin{aligned} I_1 &= R_1 \cos \theta_1, & I_3 &= R_2 \cos \theta_2, \\ I_2 &= R_1 \sin \theta_1, & I_4 &= R_2 \sin \theta_2 \end{aligned}$$

и усредним результирующую плотность вероятностей по θ_1 и θ_2 , чтобы получить вероятность того, что R_1 и R_2 лежат в интервалах dR_1 и dR_2 . Она равна

$$\frac{R_1 R_2 dR_1 dR_2}{4\pi^2 A} \int_0^{2\pi} d\theta_1 \int_0^{2\pi} d\theta_2 \exp \left\{ -\frac{1}{2A} \left[\psi_0 R_1^2 + \psi_0 R_2^2 - \right. \right. \\ \left. \left. - 2\mu_{13} R_1 R_2 \cos(\theta_2 - \theta_1) - 2\mu_{14} R_1 R_2 \sin(\theta_2 - \theta_1) \right] \right\}$$

Так как подинтегральное выражение есть периодическая функция θ_2 , то можно интегрировать в пределах от $\theta_2 = \theta_1$ до $\theta_2 = \theta_1 + 2\pi$ вместо пределов от 0 до 2π . Это интегрирование дает I_0 — функцию Бесселя первого рода от мнимого аргумента. Результирующая плотность вероятностей для R_1 и R_2 равна

$$\frac{R_1 R_2}{A} I_0 \left[\frac{R_1 R_2}{A} (\mu_{13}^2 + \mu_{14}^2)^{1/2} \right] \exp \left[-\frac{\psi_0}{2A} (R_1^2 + R_2^2) \right], \quad (3.7-13)$$

где из (3.7-12)

$$A = \psi_0^2 - \mu_{13}^2 - \mu_{14}^2,$$

а μ_{13} и μ_{14} находятся из (3.7-11). Конечно, R_1 и R_2 всегда положительны.

Для идеального полосового фильтра с частотами среза f_a и f_b положим

$$f_m = \frac{f_b + f_a}{2}, \quad \omega(f) = \omega_0 \text{ для } f_a < f < f_b$$

и получим

$$\begin{aligned} \psi_0 &= \omega_0 (f_b - f_a), \\ \mu_{13} &= \int_{f_a}^{f_b} \omega_0 \cos 2\pi(f - f_m)\tau df = \frac{\omega_0 \sin \pi(f_b - f_a)\tau}{\pi \tau}, \end{aligned}$$

$$\mu_{14} = \int_{f_a}^{f_b} \omega_0 \sin 2\pi(f - f_m)\tau df = 0.$$

Член I_0 в (3.7-13), который устанавливает корреляцию между R_1 и R_2 , приобретает вид

$$I_0 \left(\frac{R_1 R_2}{\psi_0} \frac{\frac{\sin x}{x}}{1 - \frac{\sin^2 x}{x^2}} \right),$$

где $x = \pi(f_b - f_a)\tau$. Когда x кратно π , то R_1 и R_2 — независимые случайные переменные. Когда $x = 0$, то $R_1 = R_2$. Отсюда можно приближенно заключить, что период флуктуаций огибающей R есть время, которое требуется, чтобы x возросло от 0 до π , или $(f_b - f_a)^{-1}$. Это связано с выводом, полученным в следующем разделе и утверждающем, что ожидаемое число выбросов огибающей в 1 сек. равно 0,641 $(f_b - f_a)$.

3.8. ВЫБРОСЫ ОГИБАЮЩЕЙ

Здесь будет рассмотрено распределение выбросов R . Исследование основано на выражении [ср. с (3.6—1)]

$$- dR dt \int_{-\infty}^0 p(R, 0, R'') R'' dR'', \quad (3.8-1)$$

представляющем вероятность того, что выброс R попадает в элементарный прямоугольник $dR dt$. Плотность вероятностей для распределения трех переменных R, R', R'' есть $p(R, R', R'')$, где штрихи обозначают дифференцирование по t .

Определим $p(R, R', R'')$ из плотности вероятностей для $I_c, I'_s, I''_c, I_s, I'_c, I''_s$, которые будем обозначать x_1, x_2, \dots, x_6 . Чередование I'_s и I'_c обнаруживается в последующем. Удобно ввести обозначения

$$b_n = (2\pi)^n \int_0^\infty \omega(f) (f - f_m)^n df, \quad (3.8-2)$$

$$b_0 = \psi_0,$$

где f_m есть средняя частота полосы, т. е. частота, выбранная при определении огибающей R . Как видно, b_n аналогично производным $\psi(\tau)$ при $\tau = 0$.

Из определений (3.7—3) для I_c и I_s получим моменты второго порядка:

$$\overline{x_1^2} = \overline{I_c^2} = \psi_0 = b_0,$$

$$\overline{x_4^2} = \overline{I_s^2} = b_0,$$

$$\overline{x_2^2} = \overline{I'^2_2} = \sum_1^N \omega(f_n) \Delta f 4\pi^2 (f_n - f_m)^2 = b_2,$$

$$\overline{x_5^2} = \overline{I'^2_c} = b_2,$$

$$\overline{x_3^2} = \overline{I''^2_c} = b_4,$$

$$\begin{aligned} \overline{x_6^2} &= \overline{I_s''^2} = b_4, \\ \overline{x_1 x_2} &= \overline{I_c I_s'} = \sum_1^N \omega(f_n) \Delta f 2\pi(f_n - f_m) = b_1, \\ \overline{x_4 x_5} &= \overline{I_s I_c'} = -b_1, \\ \overline{x_1 x_3} &= \overline{I_c I_c''} = -\sum_1^N \omega(f) \Delta f 4\pi^2(f_n - f_m)^2 = -b_2, \\ \overline{x_4 x_6} &= \overline{I_s I_s''} = -b_2, \\ \overline{x_2 x_3} &= \overline{I_s' I_c''} = -b_3, \\ \overline{x_5 x_6} &= \overline{I_c' I_s''} = b_3. \end{aligned}$$

Все другие моменты второго порядка равны нулю. Поэтому матрица моментов

$$M = \begin{vmatrix} b_0 & b_1 & -b_2 & 0 & 0 & 0 \\ b_1 & b_2 & -b_3 & 0 & 0 & 0 \\ -b_2 & -b_3 & b_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & b_0 & -b_1 & -b_2 \\ 0 & 0 & 0 & -b_1 & b_2 & b_3 \\ 0 & 0 & 0 & -b_2 & b_3 & b_4 \end{vmatrix}$$

Взаимная матрица равна

$$\begin{vmatrix} B_0 & B_1 & -B_2 & 0 & 0 & 0 \\ B_1 & B_{22} & -B_3 & 0 & 0 & 0 \\ -B_2 & -B_3 & B_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & B_0 & -B_1 & -B_2 \\ 0 & 0 & 0 & -B_1 & B_{22} & B_3 \\ 0 & 0 & 0 & -B_2 & B_3 & B_4 \end{vmatrix}$$

$$\begin{aligned} B_0 &= (b_2 b_4 - b_3^2) B, & B_{22} &= (b_0 b_4 - b_2^2) B, \\ B_1 &= -(b_1 b_4 - b_2 b_3) B, & B_3 &= -(b_0 b_3 - b_1 b_2) B, \\ B_2 &= (b_1 b_3 - b_2^2) B, & B_4 &= (b_0 b_2 - b_1^2) B, \\ B &= b_0 b_2 b_4 + 2 b_1 b_2 b_3 - b_2^3 - b_0 b_3^2 - b_4 b_1^2, \end{aligned}$$

$$|M| = B^2,$$

где B есть определитель матриц третьего порядка в верхнем левом и в нижнем правом углах M .

Как и в более раннем выводе, распределение x_1, \dots, x_6 нормально во всех измерениях. Показательная функция равна $-[2|M|]^{-1}$, умноженному на

$$B_0(x_1^2 + x_4^2) + 2B_1(x_1x_2 - x_4x_5) - 2B_2(x_1x_3 + x_4x_6) + \\ + B_{22}(x_2^2 + x_5^2) - 2B_3(x_2x_3 - x_5x_6) + B_4(x_3^2 + x_6^2). \quad (3.8-4)$$

В согласии с предыдущими выводами положим

$$x_1 = I_c = R \cos \theta, \quad x_4 = I_s = R \sin \theta, \\ x_2 = I'_s = R' \sin \theta + R \cos \theta \theta', \\ x_5 = I'_c = R' \cos \theta - R \sin \theta \theta', \\ x_3 = I''_c = R'' \cos \theta - 2R' \sin \theta \theta' - R \cos \theta \theta'^2 - R \sin \theta \theta'', \\ x_6 = I''_s = R'' \sin \theta + 2R' \cos \theta \theta' - R \sin \theta \theta'^2 + R \cos \theta \theta''.$$

Угол θ изменяется от 0 до 2π , а θ' и θ'' изменяются от $-\infty$ до $+\infty$. Образуюя якобиан, можно показать, что

$$dx_1 dx_2 \dots dx_6 = R^3 dR dR' dR'' d\theta d\theta' d\theta''.$$

Величины, входящие в (3.8—4), равны

$$x_1^2 + x_4^2 = R^2, \\ x_1x_2 - x_4x_5 = R^2\theta', \\ x_1x_3 + x_4x_6 = RR'' - R^2\theta'^2, \\ x_2^2 + x_5^2 = R'^2 + R^2\theta'^2, \\ x_2x_3 - x_5x_6 = RR''\theta' - 2R'^2\theta' - R'R\theta'' - R^2\theta'^3, \\ x_3^2 + x_6^2 = R''^2 - 2RR''\theta'^2 + 4R'^2\theta'^2 + 4RR'\theta'\theta'' + R^2\theta'^4 + R^2\theta''^2.$$

Выражение для $p(R, 0, R'')$ найдем, если подставим эти значения x в (3.8—4) и проинтегрируем результирующую плотность вероятностей по интервалам $\theta, \theta', \theta''$:

$$p(R, 0, R'') = \frac{R^3}{8\pi^3 B} \int_0^{2\pi} d\theta \int_{-\infty}^{+\infty} d\theta' \int_{-\infty}^{+\infty} d\theta'' \exp \left\{ -\frac{1}{2B^2} [B_0 R^2 + \right. \\ \left. + 2B_1 R^2 \theta' - 2B_2 (RR'' - R^2 \theta'^2) + B_{22} R^2 \theta'^2 - 2B_3 R \theta' (R'' - R \theta'^2) + \right. \\ \left. + B_4 (R''^2 - 2RR'' \theta'^2 + R^2 \theta'^4 + R^2 \theta''^2) \right\}. \quad (3.8-5)$$

Интегрирование по θ и θ'' может быть выполнено сразу, и $p(R, 0, R'')$ будет состоять из некратного интеграла, с которым, к сожалению, трудно иметь дело. Поэтому положим, что спектр

$\omega(f)$ симметричен относительно средней частоты полосы f_m . Из (3.8—2) b_1 и b_3 равны нулю, а из (3.8—3) B_1 и B_3 равны нулю.

При таком предположении выражение (3.8—5) приводится к виду

$$\rho(R, 0, R'') = R^2 (2\pi)^{-3/2} B_4^{-1/2} \int_{-\infty}^{+\infty} d\theta' \exp \left\{ -\frac{1}{2B^2} [B_0 R^2 + (B_{22} + 2B_2) R^2 \theta'^2 - 2B_2 R R'' + B_4 (R'' - R \theta'^2)^2] \right\}. \quad (3.8-6)$$

Вероятность того, что выброс случится в элементарном прямоугольнике $dR dt$, равна из (3.8—1) $\rho(t, R) dR dt$, где

$$\rho(t, R) = - \int_{-\infty}^0 \rho(R, 0, R'') R'' dR''. \quad (3.8-7)$$

Подставим (3.8—6) в это выражение и сделаем следующую замену переменных:

$$x = \frac{B_4^{1/2}}{\sqrt{2}B} R \theta'^2, \quad y = - \frac{B_4^{1/2}}{\sqrt{2}B} R'', \\ z = - \frac{B_2}{\sqrt{2}B_4} R = \frac{b_2^2}{\sqrt{2}B_4} R, \quad (3.8-8)$$

$$b = - \frac{(B_{22} + 2B_2)}{2B b_2^2} = \left(\frac{3}{2} - \frac{b_0 b_4}{2b_2^2} \right) = \frac{1}{3} (3 - a^2),$$

$$a^2 = \frac{B_0}{2B^2} \frac{2B_4}{b_2^2} = \frac{b_0 b_4}{b_2^2}.$$

Здесь используются выражения для B , полученные из (3.8—3), при условии, что b_1 и b_3 равны нулю. Поэтому

$$\rho(t, R) = \frac{4}{b_0 b_2^2} \left(\frac{B_2}{2\pi} \right)^{3/2} \int_0^{\infty} y dy \int_0^{\infty} x^{-1/2} dx \times \\ \times \exp [-a^2 z^2 + 2bzx + 2zy - (x+y)^2]. \quad (3.8-9)$$

Как и следовало ожидать, это выражение показывает, что $\rho(t, R)$ не зависит от t .

Ряд для $\rho(t, R)$ может быть получен путем разложения $\exp 2z(y+bx)$ и последующего почленного интегрирования. Воспользуемся выражением

$$\int_0^{\infty} dy \int_0^{\infty} dx x^\mu y^\gamma e^{-(x+y)^2} = \frac{\sqrt{\pi}}{2^{\mu+\gamma+2}} \frac{\Gamma(\gamma+1) \Gamma(\mu+1)}{\Gamma\left(\frac{\mu+\gamma+3}{2}\right)},$$

которое может быть вычислено, подставляя

$$x = \rho^2 \cos^2 \varphi, \quad y = \rho^2 \sin^2 \varphi.$$

Двойной интеграл в формуле (3.8—9) превращается в

$$\begin{aligned} e^{-a^2 z^2} \sqrt{\frac{\pi}{2}} \sum_{n=0}^{\infty} \frac{(2z)^n}{n!} \sum_{m=0}^n \frac{n! b^m}{m! (n-m)!} \frac{\Gamma\left(m + \frac{1}{2}\right) \Gamma(n-m+2)}{2^{n-2} \Gamma\left(\frac{n}{2} + \frac{7}{4}\right)} = \\ = \pi 2^{-5/4} \sum_{n=0}^{\infty} \frac{z^n e^{-a^2 z^2}}{\Gamma\left(\frac{n}{2} + \frac{7}{4}\right)} A_n, \end{aligned}$$

где $A_0 = 1$, а

$$\begin{aligned} A_n = \sum_{m=0}^n \frac{\left(\frac{1}{2}\right) \left(\frac{3}{2}\right) \cdots \left(m - \frac{1}{2}\right)}{m!} (n-m+1) b^m, \quad 0 < n \\ A_n \approx (n+1) (1-b)^{-1/4} - \frac{b}{2} (1-b)^{-3/4}, \quad n \text{ велико.} \end{aligned} \quad (3.8-10)$$

Член, соответствующий $m=0$ в (3.8—10), равен $n+1$.

Поэтому получим

$$\begin{aligned} p(t, R) = \frac{e^{-a^2 z^2}}{4 b_0 b_2^4} \frac{(Bz)^{3/4}}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{z^n}{\Gamma\left(\frac{n}{2} + \frac{7}{4}\right)} A_n = \\ = \frac{e^{-a^2 z^2}}{4\sqrt{\pi}} \frac{b^{1/4}}{b_0} (a^2-1)^{1/4} z^{3/4} \sum_{n=0}^{\infty} \frac{z^n A_n}{\Gamma\left(\frac{n}{2} + \frac{7}{4}\right)} \end{aligned} \quad (3.8-11)$$

Нас интересует ожидаемое число выбросов в 1 сек. N . Из аналогичного вывода для I следует, что N есть коэффициент при dt , когда (3.8—1) интегрируется по R в пределах от 0 до ∞ . Поэтому из (3.8—7) и из

$$dR = \sqrt{2B_4} b_2^{-2} dz = (2b_0 B)^{1/4} b_2^{3/4} dz = [2b_0 (a^2-1)]^{1/4} dz$$

находим

$$N = \int_0^{\infty} p(t, R) dR = \frac{(a^2-1)^2}{(\Sigma a)^{5/4}} \left(\frac{b_2}{\pi b_0}\right)^{1/4} \sum_{n=0}^{\infty} \frac{\Gamma\left(\frac{n}{2} + \frac{5}{4}\right)}{\Gamma\left(\frac{n}{2} + \frac{7}{4}\right)} \frac{A_n}{a^n}. \quad (3.8-12)$$

Уравнения (3.8—11) и (3.8—12) были выведены в предположении, что спектр $\omega(f)$ симметричен относительно f_m , т. е. что ослабление

полосового фильтра симметрично относительно средней частоты. Сделаем теперь дальнейший шаг и рассмотрим идеальный полосовой фильтр

$$\begin{aligned}\omega(f) &= \omega_0, & f_a < f < f_b, \\ \omega(f) &= 0, & \text{во всех других случаях,} \\ 2f_m &= f_a + f_b.\end{aligned}\tag{3.8-13}$$

Подставляя это в (3.8-2), получим нуль для b_1 и b_3 и, кроме того,

$$\begin{aligned}b_0 &= \omega_0(f_b - f_a) = \psi_0, \\ b_2 &= \frac{\pi^2 \omega_0}{3} (f_b - f_a)^3, \\ b_4 &= \frac{\pi^4 \omega_0}{5} (f_b - f_a)^5, \\ a^2 &= \frac{9}{5},\end{aligned}\tag{3.8-14}$$

$$b = \frac{1}{2} (3 - a^2) = \frac{3}{5},$$

$$R = [2b_0(a^2 - 1)]^{1/2} z = \left(\frac{8}{5} \psi_0\right)^{1/2} z,$$

$$\left(\frac{b_3}{\pi b_0}\right)^{1/2} = \left(-\frac{\pi}{3}\right)^{1/2} (f_b - f_a), \quad a^2 z^2 = \frac{9R^2}{8\psi_0}.$$

n	A_n	n	A_n
0	1	4	6,775
1	2,3	5	8,333
2	3,735	6	9,9002
3	5,238	7	11,4736

$$A_n \approx 1,5811 n + 0,3953$$

Из (3.8-12) находим, что ожидаемое число выбросов огибающей в 1 сек. на выходе идеального полосового фильтра равно

$$N = 0,64110 (f_b - f_a).\tag{3.8-15}$$

Распределение выбросов R для случая идеального полосового фильтра может быть найдено путем подстановки результатов (3.8-14) в выражение (3.8-11). Тогда получим

$$p(t, R) dR = \frac{dR}{\psi_0^{1/2}} \frac{(f_b - f_a)}{4} \sqrt{\frac{\pi}{3}} \left(\frac{4z}{5}\right)^{3/2} e^{-a^2 z^2} \sum_{n=0}^{\infty} \frac{z^n A_n}{\Gamma\left(\frac{n}{2} + \frac{7}{4}\right)}.$$

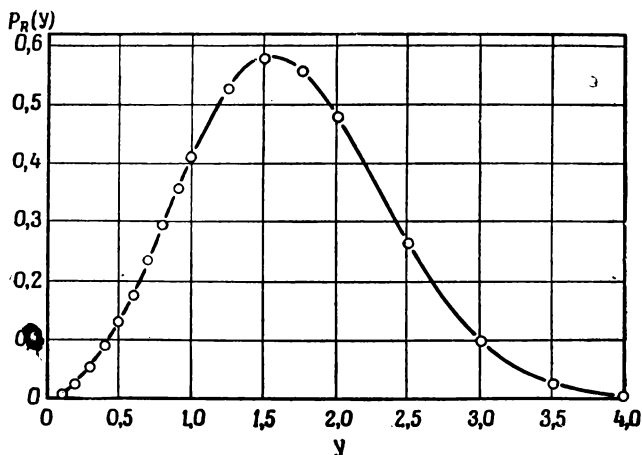
Удобно представить y как отношение

$$y = \frac{R}{\text{эфф. знач. } I(t)} = \frac{R}{\psi_0^{1/2}} = \left(\frac{8}{5}\right)^{1/2} z,$$

где подразумевается, что R соответствует выбросу огибающей. Так как значение R , соответствующее случайно выбранному выбросу огибающей, есть случайная переменная, то и y также является случайной переменной. Плотность вероятностей для y обозначим через $p_R(y)$, причем

$$p_R(y)dy = \frac{p(t, R) dR}{0,64110 (f_b - f_a)}.$$

На фиг. 3 $p_R(y)$ представлена как функция y .



Фиг. 3. Плотность вероятностей выбросов огибающей тока шумов на выходе идеального полосового фильтра.

$\frac{p_R(y)}{\sqrt{\psi_0}} dR$ - вероятность того, что случайно выбранный выброс огибающей заключен между R и $R + dR$.

Функция распределения $P(R_{\text{макс.}} < y\sqrt{\psi_0})$, определяемая как

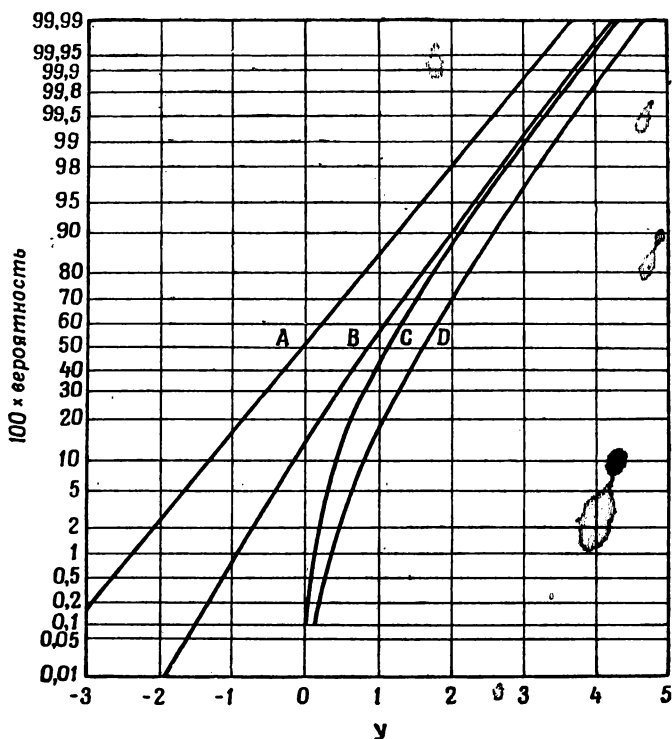
$$P(R_{\text{макс.}} < y\sqrt{\psi_0}) = \int_0^y p_R(y) dy,$$

изображена на фиг. 4 вместе с другими подобными кривыми. Эта функция представляет вероятность того, что случайно выбранный выброс огибающей будет меньше заданного значения $y\sqrt{\psi_0} = R$.

Если y велико, скажем $y > 2,5$, то

$$p_R(y) \approx \frac{\sqrt{\frac{\pi}{6}}}{0,64110} (y^2 - 1) e^{-y^2/2},$$

$$P(R_{\text{макс.}} < y \sqrt{V\psi_0}) \approx 1 - \frac{\sqrt{\frac{\pi}{6}}}{0,64110} y e^{-y^2/2}$$



Фиг. 4. Распределение вероятностей выбросов.

$A = P(I < y \sqrt{V\psi_0})$. — вероятность того, что I меньше, чем $y \sqrt{V\psi_0}$. Подобно этому $C = P(R < y \sqrt{V\psi_0})$, $B = P(I_{\text{макс.}} < y \sqrt{V\psi_0})$ — вероятность того, что произвольный выброс I меньше, чем $y \sqrt{V\psi_0}$. Подобно этому $D = P(R_{\text{макс.}} < y \sqrt{V\psi_0})$.

Приближенное выражение для $p_R(y)$ может быть получено из интеграла (3.8—9) для $p(t, R)$. Действительно, заменяя переменные интегрирования x, y в (3.8—9) на

$$x' = x, \quad y' = x + y,$$

интегрируя часть интеграла, в которую входит y' , по частям и предполагая $b < 1$ (согласно неравенству Шварца $a^2 \gg 1$, так что всегда $b \leq 1$), получим

$$\rho(t, R) \approx \left(\frac{b_2}{2\pi}\right)^{\frac{1}{2}} \frac{e^{-R^2/2\psi_0}}{\psi_0} \left(\frac{R^2}{\psi_0} - 1\right),$$

когда R — велико.

Если вместо случая идеального полосового фильтра рассмотрим случай, когда $\omega(f)$ равно¹⁾

$$\omega(f) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(f-f_m)^2/2\sigma^2}, \quad f_m \gg \sigma, \quad (3.8-16)$$

то найдем

$$b_0 = 1, \quad b_2 = 4\pi^2\sigma^2, \quad b_4 = 48\pi^4\sigma^4, \quad a^2 = 3, \quad b = 0, \quad A_n = (n+1).$$

Грубо приближенно сумма ряда в (3.8—12) равна 3,97. Поэтому ожидаемое число выбросов огибающей в 1 сек. равно

$$N = 2,52\sigma. \quad (3.8-17)$$

Полоса пропускания определяется коэффициентом σ . Представляется трудным сравнить этот случай со случаем идеального полосового фильтра. Если воспользоваться тем обстоятельством, что фильтр, для которого

$$\omega(f) = \omega_0 \exp\left[-\pi\left(\frac{f-f_m}{f_b-f_a}\right)^2\right],$$

пропускает ту же самую среднюю мощность, как и идеальный полосовой фильтр с полосой пропускания, равной $f_b - f_a$, то получим

$$f_b - f_a = \sigma\sqrt{2\pi}$$

и выражение для N превращается в

$$N = 1,006(f_b - f_a).$$

3.9. ФЛУКТУАЦИИ ЭНЕРГИИ

Здесь исследуем статистические свойства случайной переменной

$$E = \int_{t_1}^{t_1+T} I^2(t) dt, \quad (3.9-1)$$

где $I(t)$ — ток шумов, а t_1 выбрано случайно.

¹⁾ Уравнение (3.8—16) нормировано в том смысле, что средняя мощность ψ_0 принята равной единице. Здесь σ имеет другое значение, чем раньше, и размерность ее есть размерность частоты. (Прим. ред.)

Если пользоваться представлениями (2.8—1) или (2.8—6), то можно написать случайную переменную E в виде

$$E = \int_{T/2}^{T/2} I^2(t) dt, \quad (3.9-2)$$

где случайный характер выражению, стоящему справа, придается наличием либо a_n и b_n , если используется представление (2.8—1), либо φ_n , если применяется (2.8—6).

Среднее значение E обозначим через m_T , где из (3.1—2)

$$\bar{E} = m_T = \int_{-T/2}^{T/2} \overline{I^2(t)} dt = \int_{-T/2}^{T/2} \psi(0) dt = T\psi_0 = T \int_0^{\infty} w(f) df. \quad (3.9-3)$$

Момент второго порядка равен

$$\bar{E}^2 = \int_{-T/2}^{T/2} dt_1 \int_{-T/2}^{T/2} dt_2 \overline{I^2(t_1) I^2(t_2)}. \quad (3.9-4)$$

Если теперь положим $t_2 = t_1 + \tau$, то, как видно из раздела 3.2, получим выражение плотностей вероятностей для $I(t_1)$ и $I(t_1 + \tau)$ и, следовательно, можем написать для искомого среднего значения

$$\begin{aligned} \overline{I_1^2 I_2^2} &= \frac{1}{2\pi A} \int_{-\infty}^{+\infty} dI_1 \int_{-\infty}^{+\infty} dI_2 I_1^2 I_2^2 \times \\ &\times \exp \left[-\frac{1}{2A^2} \left(\psi_0 I_1^2 + \psi_0 I_2^2 - 2\psi_\tau I_1 I_2 \right) \right] \quad (3.9-5) \\ A^2 &= \psi_0^2 - \psi_\tau^2, \quad I_1 = I(t_1), \quad I_2 = I(t_1 + \tau) = I(t_2). \end{aligned}$$

Интеграл может быть вычислен при помощи (3.5—6), если положить

$$\begin{aligned} I_1 &= Ax \sqrt{\frac{2}{\psi_0}}, \quad I_2 = Ay \sqrt{\frac{2}{\psi_0}}, \\ \psi_\tau &= -\psi_0 \cos \varphi, \quad A = \psi_0 \sin \varphi. \quad (3.9-6) \end{aligned}$$

Поэтому

$$\overline{I_1^2 I_2^2} = \psi_0^2 (1 + 2 \cos^2 \varphi) = \psi_0^2 + 2\psi_\tau^2. \quad (3.9-7)$$

Кстати, это дает выражение функции корреляций для $I^2(t)$. Заменяя τ его значением $t_2 - t_1$ и возвращаясь к (3.9—4), получим

$$\bar{E}^2 = T^2 \psi_0^2 + 2 \int_{-T/2}^{T/2} dt_1 \int_{-T/2}^{T/2} dt_2 \psi^2(t_2 - t_1). \quad (3.9-8)$$

Если введем σ_T — стандартное отклонение E

$$\sigma_T^2 = \overline{E^2} - m_T^2,$$

то найдем

$$\begin{aligned} \sigma_T^2 &= \overline{(E - \bar{E})^2} = 2 \int_{-T/2}^{T/2} dt_1 \int_{-T/2}^{T/2} dt_2 \psi^2(t_2 - t_1) = \\ &= 4 \int_0^T (T - x) \psi^2(x) dx, \end{aligned}$$

где вторая строчка может быть получена из первой либо путем замены переменных интегриации, как в (3.9—27), либо методом, применяемым ниже при рассмотрении \bar{E}^3 . Следует заметить, что пределы интегрирования — $T/2$, $T/2$ в двойном интеграле могут быть заменены на 0, T путем замены переменной $t = t' - T/2$ как для t_1 , так и для t_2 .

Если воспользуемся выражением

$$\psi(\tau) = \int_0^{\infty} \omega(f) \cos 2\pi f \tau \, df, \quad (2.1-6)$$

то получим

$$\begin{aligned} \sigma_T^2 &= \int_0^{\infty} \omega(f_1) \, df_1 \int_0^{\infty} \omega(f_2) \, df_2 \left[\frac{\sin^2 \pi(f_1 + f_2) T}{\pi^2 (f_1 + f_2)^2} + \right. \\ &\quad \left. + \frac{\sin^2 \pi(f_1 - f_2) T}{\pi^2 (f_1 - f_2)^2} \right] \end{aligned} \quad (3.9-9)$$

Если эту формулу применить к случаю сравнительно узкополосного полосового фильтра и если $T(f_b - f_a) \gg 1$, то членом с $f_1 + f_2$ можно будет пренебречь и получить приближение

$$\begin{aligned} \sigma_T^2 &\approx \int_{f_a}^{f_b} \omega_0^2 df_1 \int_{-\infty}^{+\infty} \omega_0 df_2 \frac{\sin^2 \pi(f_1 - f_2) T}{\pi^2 (f_1 - f_2)^2} = \\ &= \omega_0^2 T (f_b - f_a) = \omega_0 m_T, \end{aligned} \quad (3.9-10)$$

где из (3.9—3)

$$m_T = \omega_0 T (f_b - f_a). \quad (3.9-11)$$

Момент третьего порядка $\overline{E^3}$ может быть вычислен подобным же образом. Однако в этом случае имеет смысл ввести характеристическую функцию для распределения $I(t_1)$, $I(t_2)$, $I(t_3)$. Так

как это распределение нормально, то его характеристическая функция равна

$$\begin{aligned} & \text{сред. exp } (iz_1 I_1 + iz_2 I_2 + iz_3 I_3) = \\ & = \exp \left\{ - \left[\frac{\psi_0}{2} (z_1^2 + z_2^2 + z_3^2) + \psi(t_2 - t_1) z_1 z_2 + \right. \right. \\ & \quad \left. \left. + \psi(t_3 - t_1) z_1 z_3 + \psi(t_3 - t_2) z_2 z_3 \right] \right\} \end{aligned} \quad (3.9-12)$$

Из определения характеристической функции следует

$$\begin{aligned} \overline{I_1^2 I_2^2 I_3^2} &= - \left[\text{коэф. при } \frac{z^2 z_2^2 z_3^2}{2! 2! 2!} \text{ в х. ф.} \right] = \\ &= \psi_0^3 + 2\psi_0 (\psi_{21}^2 + \psi_{31}^2 + \psi_{32}^2) + 8\psi_{21} \psi_{31} \psi_{32}, \end{aligned} \quad (3.9-13)$$

где $\psi(t_2 - t_1), \dots$ обозначены, соответственно, через ψ_{21} и т. д. Если (3.9-13) умножить на $dt_1 dt_2 dt_3$, затем проинтегрировать в пределах от 0 до T и воспользоваться приведенным выше двойным интегралом для σ_T^2 , то получим

$$\overline{(E - \bar{E})^3} = 2! 2^2 \int_0^T dt_1 \int_0^T dt_2 \int_0^T dt_3 \psi_{21} \psi_{31} \psi_{32}.$$

Обозначая тройной интеграл справа через J и дифференцируя, имеем

$$\begin{aligned} \frac{dJ}{dT} &= 3 \int_0^T dt_1 \int_0^T dt_2 \psi(t_2 - t_1) \psi(T - t_1) \psi(T - t_2) = \\ &= 3 \int_0^T dx \int_0^T dy \psi(x - y) \psi(x) \psi(y) = 6 \int_0^T dx \int_0^x dy \psi(x - y) \psi(x) \psi(y). \end{aligned}$$

При переходе от первой строчки ко второй t_1 и t_2 были заменены на $T - x$ и $T - y$ соответственно. При переходе от второго выражения к третьему использовались соотношения, символически представленные в виде

$$\int_0^T dx \int_0^T dy = \int_0^T dx \int_0^x dy + \int_0^T dx \int_x^T dy = \int_0^T dx \int_0^x dy + \int_0^T dy \int_0^y dx,$$

а также то обстоятельство, что подинтегральное выражение симметрично по x и y . Интегрируя dJ/dT по T в пределах от 0 до T_1 , пользуясь формулой

$$\int_0^{T_1} dT \int_0^T f(x) dx = \int_0^{T_1} (T_1 - x) f(x) dx,$$

замечая, что $J=0$, когда $T=0$, и, наконец, опуская индекс при T_1 , окончательно получим

$$\overline{(E - \bar{E})^3} = 48 \int_0^T dx \int_0^x dy (T - x) \psi(x) \psi(y) \psi(x - y).$$

Аналогично можно свести к сумме двух тройных интегралов и \bar{E}^4 . До сих пор в этом разделе говорилось о статистических константах E . Нахождение точного выражения для плотности вероятностей E , в котором T рассматривается как параметр, представляется весьма трудным.

Когда T очень мало, то E приближенно равно $I^2(t)T$. Вероятность того, что E заключено в интервале dE , равна вероятности нахождения тока в промежутке $(-I, -I-dI)$ плюс вероятность нахождения тока в промежутке $(I, I+dI)$:

$$\frac{2dI}{\sqrt{2\pi\psi_0}} \exp\left(-\frac{I^2}{2\psi_0}\right) = \frac{1}{\sqrt{2\pi\psi_0 ET}} \exp\left(-\frac{E}{2\psi_0 T}\right) dE, \tag{3.9-14}$$

где E положительно,

$$I = \sqrt{\frac{E}{T}}, \quad dI = \frac{1}{2\sqrt{ET}} dE,$$

а T предполагается настолько малым, что в течение интервала длительностью T $I(t)$ значительно не изменяется.

Когда T очень велико, можно разделить его на ряд интервалов, скажем n , длительностью каждый T/n . Пусть E_r будет доля энергии от r -го интервала. Тогда энергия в полном интервале

$$E = E_1 + E_2 + \dots + E_n.$$

Если частичные интервалы достаточно велики, то E_r — существенно независимые случайные переменные. Если, в дополнение, n достаточно велико, то E приближенно распределено по нормальному закону. Следовательно, когда T весьма велико, вероятность того, что E заключено в интервале dE , равна

$$\frac{dE}{\sigma_T \sqrt{2\pi}} \exp\left[-\frac{(E - m_T)^2}{2\sigma_T^2}\right] \tag{3.9-15}$$

где

$$m_T = T \int_0^\infty \omega(f) df, \tag{3.9-16}$$

$$\sigma_T^2 = T \int_0^\infty \omega^2(f) df,$$

причем второе соотношение получено, полагая в (3.9—9) $T \rightarrow \infty$. Аналогия с теоремой о наложении случайных возмущений (раздел 1.2) очевидна.

Если имеем дело с полосовым фильтром, то можно пользоваться (3.9—10) и (3.9—11).

Рассмотрим полосовой фильтр со сравнительно узкой полосой пропускания, так что можно найти T , для которого $Tf_a \gg 2\pi$, но $T(f_b - f_a) \ll 0,64$. Следовательно, в T содержится несколько периодов частоты f_a , но, согласно (3.8—15), в течение этого интервала огибающая значительно не изменяется. Поэтому на протяжении этого интервала $I(t)$ может рассматриваться как синусоидальная волна с амплитудой R . Соответствующее значение E приближенно равно

$$E = T \frac{R^2}{2},$$

где распределение огибающей R берется из (3.7—10). Отсюда следует, что вероятность нахождения E в интервале dE равна

$$\frac{dE}{\psi_0 T} \exp\left(-\frac{E}{\psi_0 T}\right) = \frac{dE}{m_T} e^{-E/m_T}, \quad (3.9-17)$$

когда E мало, но не слишком.

Рассматривая (3.9—14) и (3.9—17), замечаем, что они имеют вид

$$\frac{a^{n+1} E^n}{\Gamma(n+1)} e^{-aE} dE. \quad (3.9-18)$$

Кроме того, нормальный закон (3.9—15) может быть отсюда получен, полагая, что n становится большим.

Этим подразумевается, что приближенное выражение для распределения E дается формулой (3.9—18), если a и n выбраны так, что значения m_T и σ_T получаются из (3.9—3) и (3.9—9). Это дает

$$a = \frac{m_T}{\sigma_T^2} \quad n + 1 = \frac{m_T^2}{\sigma_T^2} \quad (3.9-19)$$

и если опустим индекс T и подставим значение a в (3.9—18), то получим

$$\frac{\left(\frac{mE}{\sigma^2}\right)^n}{\Gamma(n+1)} \exp\left(-\frac{mE}{\sigma^2}\right) d\left(\frac{mE}{\sigma^2}\right) \quad n = \frac{m^2}{\sigma^2} - 1. \quad (3.9-20)$$

Представление об этом распределении дает следующая таблица:

n	$T(f_b - f_a)$	$x_{0,25}$	$x_{0,50}$	$x_{0,75}$	$\frac{x_{0,25}}{x_{0,50}}$	$\frac{x_{0,75}}{x_{0,50}}$
0	0	0,29	0,695	1,39	0,415	2,00
1	1,45	0,96	1,68	2,69	0,572	1,60
2	2,4	1,73	2,67	3,94	0,647	1,47
3	3,4	2,54	3,67	5,12	0,692	1,39
5	5,4	4,22	5,67	7,42	0,744	1,31
10	10,5	8,63	10,67	13,02	0,808	1,22
24	25	21,47	24,67	28,17	0,870	1,14
48	50	44,1	48,7	53,5	0,905	1,10

где n — показатель степени, входящий в (3.9—20). Столбец $T(f_b - f_a)$ справедлив только для полосового фильтра с узкой полосой пропускания. Цифры в этом столбце не очень точны. Следующие три столбца дают точки, которые делят распределение на четыре интервала равной вероятности

$$x_{0,25} = \frac{mE_{0,25}}{\sigma^2}, \quad E_{0,25} = \text{энергия, превышаемая в течение 75\% времени}$$

$$x_{0,50} = \frac{mF_{0,50}}{\sigma^2}, \quad E_{0,50} = \text{-----} \text{-----} 50\% \text{-----} \text{-----}$$

$$x_{0,75} = \frac{mE_{0,75}}{\sigma^2}, \quad E_{0,75} = \text{-----} \text{-----} 25\% \text{-----} \text{-----}$$

Цифры в этих столбцах получены из таблицы неполных гамма-функций Пирсона. Последние два столбца показывают, как распределение группируется вокруг среднего значения, когда оно сходится к нормальному закону.

При больших значениях n достигается нормальный закон (3.9—15). Так как в случае справедливости этого закона точки 25, 50 и 75% соответствуют $m - 0,675\sigma$, m и $m + 0,675\sigma$, то в первом приближении

$$x_{0,50} = \frac{m^2}{\sigma^2} = (n+1) \approx T(f_b - f_a),$$

$$x_{0,25} = \frac{m}{\sigma^2} (m - 0,675\sigma) = x_{0,50} - 0,675\sqrt{x_{0,50}}, \quad (3.9-21)$$

$$x_{0,75} = x_{0,50} + 0,675\sqrt{x_{0,50}}.$$

Это согласуется с таблицей (см. также фиг. 5).

Исследуем теперь флуктуации интеграла

$$A(t) = \int_{-\infty}^t I^2(\tau) e^{-\alpha(t-\tau)} d\tau. \quad (3.9-22)$$

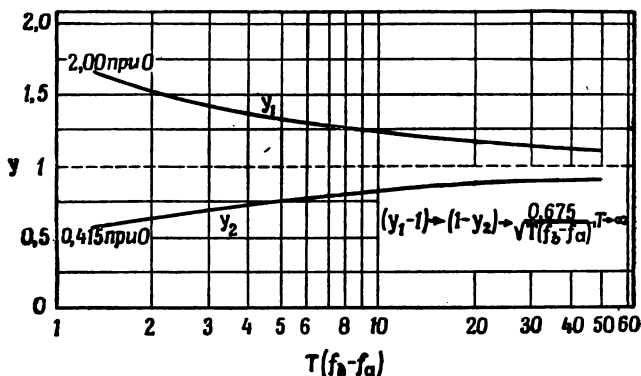
Показания термоэлектрического амперметра, через который проходит ток I , пропорциональны $A(t)$. α — постоянная прибора.

Начнем исследование $A(t)$ с нахождения его функции корреляции. Сплошной участок энергетического спектра $A(t)$ дается урав-

нением (3.9—30). Если энергетический спектр $I(f)$, обозначаемый $\omega(f)$, равен всюду нулю за исключением участка $f_a < f < f_b$, где он равен ω_0 , то энергетический спектр $A(t)$ равен

$$\frac{2\omega_0^2(f_b - f_a - f)}{a^2 + 4\pi^2 f^2} \quad \text{для } 0 < f < f_b - f_a$$

и равен нулю от $f_b - f_a$ до $2f_a$. Спектр на участке от $2f_a$ до $2f_b$ не равен нулю и может быть найден из (3.9—34). Средний квадрат флук-



Фиг. 5. Шумы теплового движения на выходе фильтра — разброс энергии флуктуаций.

$$E = \int_{t_1}^{t_1+T} I^2(t) dt, \quad t_1 - \text{взято произвольно, } I - \text{ток шумов.}$$

$$y_1 = \frac{E_{0,75}}{E_{0,50}} \quad y_2 = \frac{E_{0,25}}{E_{0,50}} \quad f_b - f_a - \text{ширина полосы фильтра.}$$

туаций $A(t)$ дается в общем случае уравнениями (3.9—28) и (3.9—32). Для полосового фильтра, если $\frac{f_b - f_a}{a}$ велико,

$$\text{эфф. знач. } \left[\frac{A(t) - \bar{A}}{\bar{A}} \right] = \left[\frac{a}{2(f_b - f_a)} \right]^{1/2}$$

Начнем с того, что положим $\tau = t - u$, благодаря чему интеграл для $A(t)$ преобразуется в

$$A(t) = \int_0^{\infty} I^2(t - u) e^{-au} du. \quad (3.9-23)$$

Чтобы получить функцию корреляции $\Psi(\tau)$ для $A(t)$, умножим $A(t)$ на $A(t + \tau)$ и усредним по всем возможным значениям токов

$$\Psi(\tau) = \overline{A(t) A(t + \tau)} = \int_0^{\infty} e^{-au} du \int_0^{\infty} e^{-av} dv \overline{I^2(t - u) I^2(t + \tau - v)}$$

Так же как и в (3.9—4), среднее в подинтегральном выражении есть функция корреляции для $I^2(t)$, аргумент равен $t + \tau - v - t + u = \tau + u - v$. Из (3.9—7) видно, что она равна

$$\psi_0^2 + 2\psi^2(\tau + u - v),$$

где $\psi(\tau)$ есть функция корреляции для $I(t)$. Следовательно,

$$\Psi(\tau) = \frac{\psi_0^2}{\alpha^2} + 2 \int_0^\infty du \int_0^\infty dv e^{-\alpha u - \alpha v} \psi^2(\tau + u - v). \quad (3.9—24)$$

Из интеграла (3.9—23) для $A(t)$ видно, что среднее значение $A(t)$ равно

$$\bar{A} = \frac{\bar{I}^2}{\alpha} = \frac{\psi_0}{\alpha}, \quad (3.9—25)$$

где

$$\psi_0 = \psi(0) = \int_0^\infty \omega(f) df = \bar{I}^2.$$

Применяя последнее соотношение опять, но только на этот раз к $A(t)$, получим

$$\overline{A^2(t)} = \Psi(0) = \bar{A}^2 + 2 \int_0^\infty du \int_0^\infty dv e^{-\alpha u - \alpha v} \psi^2(u - v). \quad (3.9—26)$$

Двойные интегралы могут быть преобразованы путем замены переменных $u + v = x$, $u - v = y$. Тогда (3.9—24) превращается в

$$\begin{aligned} \Psi(\tau) &= \bar{A}^2 + \left(\int_0^\infty dy \int_y^\infty dx + \int_{-\infty}^0 dy \int_{-y}^0 dx \right) e^{-\alpha x} \psi^2(\tau + y) = \\ &= \bar{A}^2 + \frac{1}{\alpha} \int_0^\infty e^{-\alpha y} [\psi^2(\tau + y) + \psi^2(\tau - y)] dy. \end{aligned} \quad (3.9—27)$$

Если воспользоваться тем обстоятельством, что $\psi(y)$ есть четная функция y , то из (3.9—26) следует, что средний квадрат флуктуаций $A(t)$ равен

$$\overline{[A(t) - \bar{A}]^2} = \overline{A^2(t)} - \bar{A}^2 = \frac{2}{\alpha} \int_0^\infty e^{-\alpha y} \psi^2(y) dy. \quad (3.9—28)$$

Функция $\Psi(\tau)$ может быть записана при помощи интегралов, в которые входит $\omega(f)$ — энергетический спектр $I(t)$. Вывод начинается с (3.9—24); он аналогичен переходу от (3.9—8) к (3.9—9).

В результате получим

$$\Psi(\tau) = \bar{A}^2 + \int_0^{\infty} df_1 \int_0^{\infty} df_2 \omega(f_1) \omega(f_2) \left\{ \frac{\cos 2\pi(f_1+f_2)\tau}{\alpha^2 + [2\pi(f_1+f_2)]^2} + \frac{\cos 2\pi(f_1-f_2)\tau}{\alpha^2 + [2\pi(f_1-f_2)]^2} \right\}.$$

Удобно обозначить через $\omega(-f)$ спектр для отрицательных частот, равный $\omega(f)$. Интегрирование по f_2 может быть тогда произведено в пределах от $-\infty$ до $+\infty$; в итоге получим

$$\Psi(t) = \bar{A}^2 + \int_0^{\infty} df_1 \int_{-\infty}^{+\infty} df_2 \omega(f_1) \omega(f_2) \frac{\cos 2\pi(f_1-f_2)\tau}{\alpha^2 + [2\pi(f_1-f_2)]^2}. \quad (3.9-29)$$

Энергетический спектр $W(f)$ для $A(t)$ можно найти, интегрируя $\Psi(\tau)$,

$$W(f) = 4 \int_0^{\infty} \Psi(\tau) \cos 2\pi f \tau \, d\tau.$$

Рассмотрим часть $A(t)$, подверженную флуктуациям, т. е. $[A(t) - \bar{A}]$. Ее энергетический спектр $W_c(f)$ равен

$$W_c(f) = 4 \int_0^{\infty} [\Psi(\tau) - \bar{A}^2] \cos 2\pi f \tau \, d\tau.$$

Интегрирование упрощается при применении формулы интеграла Фурье в виде

$$\int_0^{\infty} d\tau \int_{-\infty}^{+\infty} df_2 F(f_2) \cos 2\pi(u-f_2)\tau = \frac{1}{2} F(u).$$

Получим

$$\begin{aligned} W_c(f) &= \frac{1}{\alpha^2 + 4\pi^2 f^2} \int_0^{\infty} df_1 [\omega(f_1) \omega(f+f_1) + \omega(f_1) \omega(-f+f_1)] = \\ &= \frac{1}{\alpha^2 + 4\pi^2 f^2} \int_{-\infty}^{+\infty} \omega(f_1) \omega(f-f_1) df_1. \end{aligned} \quad (3.9-30)$$

Простота этого результата наводит на мысль, что может быть найден менее сложный вывод. Если попытаться воспользоваться формулой

$$\bar{\omega}(f) = \lim_{T \rightarrow \infty} \frac{2|S(f)|^2}{T}, \quad (2.5-3)$$

где $S(f)$ берется из (2.1-2), то нужно доказать, что

$$\lim_{T \rightarrow \infty} \frac{2}{T} \int_0^T dt_1 \int_0^T dt_2 e^{2\pi i f(t_1 - t_2)} I^2(t_1) I^2(t_2) =$$

$$= \int_{-\infty}^{+\infty} \omega(f_1) \omega(f - f_1) df_1, \quad (3.9-31)$$

где $f > 0$, $I(t)$ есть ток шумов, а $\omega(f)$ — его энергетический спектр. Это можно доказать, пользуясь (3.9-7) и соотношением

$$8 \int_0^{\infty} \psi^2(\tau) \cos 2\pi f \tau d\tau = \int_{-\infty}^{+\infty} \omega(x) \omega(f-x) dx$$

в соответствии с уравнением (4С-6) в Приложении 4С.

Выражение для среднего квадрата флуктуаций $A(t)$, в которое входило бы $\omega(f)$, может быть найдено, полагая в (3.9-29) $\tau=0$:

$$\overline{(A(t) - \bar{A})^2} = \Psi(0) - \bar{A}^2 = \int_0^{\infty} df_1 \int_{-\infty}^{+\infty} df_2 \frac{\omega(f_1) \omega(f_2)}{\alpha^2 + 4\pi^2(f_1 - f_2)^2}. \quad (3.9-32)$$

Тот же самый результат получим, интегрируя $W_c(f)$ из (3.9-30) в пределах от 0 до ∞ :

$$\int_0^{\infty} \frac{df}{\alpha^2 + 4\pi^2 f^2} \int_{-\infty}^{+\infty} df_1 \omega(f_1) \omega(f - f_1). \quad (3.9-33)$$

Хотя по внешнему виду это выражение и отличается от (3.9-32), но оно может быть представлено в такой же форме при помощи соотношения $\omega(-f) = \omega(f)$.

Предположим, что $I(t)$ — ток через идеальный полосовой фильтр, так что $\omega(f) = 0$ за исключением полосы $f_a < f < f_b$, где $\omega(f) = \omega_0$. Тогда, если $3f_a > f_b$,

$$\bar{A} = \frac{\omega_0}{\alpha} (f_b - f_a),$$

$$\int_{-\infty}^{+\infty} \omega(x) \omega(f-x) dx = \begin{cases} 2\omega_0^2 (f_b - f_a - f), & 0 < f < f_b - f_a, \\ \omega_0^2 (f - 2f_a), & 2f_a < f < f_b + f_a, \\ \omega_0^2 (2f_b - f), & f_b + f_a < f < 2f_b \end{cases} \quad (3.9-34)$$

и равно нулю за пределами этих диапазонов. Энергетический спектр $W_c(f)$ немедленно может быть найден по (3.9-30) путем деления этих значений на $\alpha^2 + 4\pi^2 f^2$.

Из (3.9-33)

$$\begin{aligned} \overline{(A(t) - \bar{A})^2} &= 2\omega_0^2 \int_0^{f_b - f_a} \frac{f_b - f_a - f}{\alpha^2 + 4\pi^2 f^2} df + \\ &+ \omega_0^2 \int_{2f_a}^{f_b + f_a} \frac{(f - 2f_a)}{\alpha^2 + 4\pi^2 f^2} df + \omega_0^2 \int_{f_b + f_a}^{2f_b} \frac{(2f_b - f)}{\alpha^2 + 4\pi^2 f^2} df. \end{aligned}$$

Если желательно получить точное решение, то интегрирование должно быть выполнено. Положив, что $f_b - f_a \ll f_b + f_a$, можно получить приближения для последних двух интегралов

$$\begin{aligned} \overline{(A(t) - \bar{A})^2} \approx \omega_0^2 \left[\frac{f_b - f_a}{\pi a} \operatorname{arctg} \frac{2\pi(f_b - f_a)}{a} - \right. \\ \left. - \frac{1}{4\pi^2} \log \frac{a^2 + 4\pi^2(f_b - f_a)^2}{a^2} + \frac{(f_b - f_a)^2}{a^2 + 4\pi^2(f_b + f_a)^2} \right]. \end{aligned}$$

Далее, если $\frac{2\pi(f_b - f_a)}{a}$ велико, то

$$\overline{(A(t) - \bar{A})^2} \approx \omega_0^2 \frac{f_b - f_a}{2a}$$

и относительное эффективное значение флуктуаций равно

$$\text{эфф. знач.} \left[\frac{A(t) - \bar{A}}{\bar{A}} \right] \approx \left[\frac{a}{2(f_b - f_a)} \right]^{1/2}$$

Этот результат может быть также получен из (3.9—10) и (3.9—11), полагая a настолько малым, что интеграл для $A(t)$ может быть разбит на большое число интегралов, каждый из которых охватывает интервал длительностью T . Предполагается, что aT настолько мало, что e^{-au} существенно постоянно в каждом интервале.

§3.10. РАСПРЕДЕЛЕНИЕ СУММЫ ИЗ ТОКА ШУМОВ И СИНУСОИДАЛЬНОГО ТОКА

Допустим, что имеется установившийся синусоидальный ток

$$I_p = I_p(t) = P \cos(\omega_p t - \varphi_p). \quad (3.10-1)$$

Выберем случайные моменты времени t_1, t_2, \dots и заметим соответствующие значения тока. Как распределены эти значения? Выбор случайных моментов времени в (3.10—1) со статистической точки зрения представляет собой то же самое, что и сохранение t постоянным и выбор случайных фазовых углов φ_p в диапазоне от 0 до 2π . Если I_p будет рассматриваться как случайная переменная, определяемая случайной переменной φ_p , то ее характеристическая функция равна

$$\text{сред. } e^{izI_p} = \frac{1}{2\pi} \int_0^{2\pi} e^{izP \cos(\omega_p t - \varphi)} d\varphi = J_0(Pz), \quad (3.10-2)$$

где $J_0(Pz)$ — функция Бесселя. Плотность вероятностей для I_p

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-izI_p} J_0(Pz) dz = \begin{cases} \frac{1}{\pi} (P^2 - I_p^2)^{-1/2}, & |I_p| < P, \\ 0 & |I_p| > P. \end{cases} \quad (3.10-3)$$

В этом случае проще найти плотность вероятностей непосредственно из (3.10—1), а не при помощи характеристической функции.

Допустим теперь, что имеется ток шумов I_N плюс синусоидальный ток. Сопоставляя выражение для I_N (2.8—6) с изложенным выше понятием о случайных фазовых углах φ_p , приходим к следующему представлению:

$$I(t) = I = I_p + I_N = P \cos(\omega_p t - \varphi_p) + \sum_1^M c_n \cos(\omega_n t - \varphi_n), \quad (3.10-4)$$

$$c_n^2 = 2\omega(f_n) \Delta f,$$

где φ_p и $\varphi_1, \dots, \varphi_M$ — независимые случайные углы.

Если наблюдать I в случайные моменты времени t_1, t_2, \dots , то как будут распределены отмеченные величины? Так как I_p и I_N могут рассматриваться как независимые случайные переменные и так как характеристическая функция суммы двух таких переменных равна произведению их характеристических функций, то из (3.1—6) и (3.10—2) имеем

$$\text{сред. } e^{izI} = \text{сред. } e^{iz(I_p + I_N)} = J_0(Pz) \exp\left(\frac{-\psi_0 z^2}{2}\right). \quad (3.10-5)$$

Это выражение представляет собой характеристическую функцию для I . Плотность вероятностей I равна¹⁾

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp\left[-izI - \left(\frac{\psi_0 z^2}{2}\right)\right] J_0(Pz) dz =$$

$$= \frac{1}{\pi \sqrt{2\pi\psi_0}} \int_0^\pi \exp\left[\frac{-(I - P \cos \theta)^2}{2\psi_0}\right] d\theta. \quad (3.10-6)$$

Подобным же образом можно показать, что плотность вероятностей для (I_1, I_2) , где $I_1 = I(t)$ — ток шумов плюс синусоидальный ток (3.10—4), а $I_2 = I(t + \tau)$ — значение этой суммы спустя некоторый постоянный промежуток времени τ , равна

$$\frac{(\psi_0^2 - \psi_\tau^2)^{-1/2}}{2\pi} \int_0^{2\pi} d\theta \exp\left[-\frac{B(\theta)}{2(\psi_0^2 - \psi_\tau^2)}\right], \quad (3.10-7)$$

где

$$B(\theta) = \psi_0 \{(I_1 - P \cos \theta)^2 + [I_2 - P \cos(\theta + \omega_p \tau)]^2\} - 2\psi_\tau (I_1 - P \cos \theta) [I_2 - P \cos(\theta + \omega_p \tau)].$$

¹⁾ Другой вывод этого выражения дан У. Р. Беннетом, *BSTJ*, 97, январь 1944.

Характеристическая функция для I_1 и I_2 равна

$$\begin{aligned} \text{сред. } e^{iuI_1 + ivI_2} &= J_0(P \sqrt{u^2 + v^2 + 2uv \cos \omega_p \tau}) \times \\ &\times \exp \left[-\frac{\psi_0}{2} (u^2 + v^2) - \psi_\tau uv \right]. \end{aligned} \quad (3.10-8)$$

Иногда представляет интерес распределение огибающей тока

$$I = P \cos pt + I_N. \quad (3.10-9)$$

Здесь ω_p заменено на p , а φ_p положено равным нулю. Под огибающей подразумевается $R(t)$, которую можно найти из

$$R^2(t) = R^2 = (P + I_c)^2 + I_s^2, \quad (3.10-10)$$

где I_c есть составляющая I_N «в фазе» с $\cos pt$, а I_s — составляющая «в фазе» с $\sin pt$:

$$\begin{aligned} I_c &= \sum c_n \cos [(\omega_n - p)t - \varphi_n], \\ I_s &= \sum c_n \sin [(\omega_n - p)t - \varphi_n], \\ I_N &= I_c \cos pt - I_s \sin pt, \\ \overline{I_N^2} &= \overline{I_c^2} = \overline{I_s^2} = \psi_0. \end{aligned}$$

Так как I_c и I_s распределены по нормальному закону вокруг нуля с дисперсией ψ_0 , плотности вероятностей переменных

$$x = P + I_c, \quad y = I_s$$

равны соответственно

$$\begin{aligned} (2\pi\psi_0)^{-1/2} \exp \left[-\frac{(x-P)^2}{2\psi_0} \right], \\ (2\pi\psi_0)^{-1/2} \exp \left(-\frac{y^2}{2\psi_0} \right). \end{aligned}$$

Положив

$$x = R \cos \theta, \quad y = R \sin \theta$$

и пользуясь этими распределениями, видим, что вероятность нахождения точки (x, y) в кольце $(R, R+dR)$ равна

$$\begin{aligned} \frac{R dR}{2\pi\psi_0} \int_0^{2\pi} \exp \left[-\frac{1}{2\psi_0} (R^2 + P^2 - 2RP \cos \theta) \right] d\theta = \\ = \frac{R dR}{\psi_0} \exp \left(-\frac{R^2 + P^2}{2\psi_0} \right) I_0 \left(\frac{RP}{\psi_0} \right), \end{aligned} \quad (3.10-11)$$

где I_0 — функция Бесселя мнимого аргумента

$$I_0(z) = \sum_{n=0}^{\infty} \frac{z^{2n}}{2^{2n} n! n!},$$

которая является табличной функцией. Из (3.10—11), отбрасывая dR , получаем плотность вероятностей огибающей R .

Среднее значение R^n может быть найдено путем умножения (3.10—11) на R^n и интегрирования в пределах от 0 до ∞ . Разложение функции Бесселя и почленное интегрирование дают

$$\begin{aligned} \overline{R^n} &= (2\psi_0)^{n/2} \Gamma\left(\frac{n}{2} + 1\right) e^{-P^2/2\psi_0} {}_1F_1\left(\frac{n}{2} + 1; 1; \frac{P^2}{2\psi_0}\right) = \\ &= (2\psi_0)^{n/2} \Gamma\left(\frac{n}{2} + 1\right) {}_1F_1\left(-\frac{n}{2}; 1; -\frac{P^2}{2\psi_0}\right), \end{aligned} \quad (3.10-12)$$

где ${}_1F_1$ есть гипергеометрическая функция¹⁾. При переходе от первой строчки ко второй было использовано первое преобразование Куммера для этой функции. В частном случае

$$\overline{R^2} = P^2 + 2\psi_0. \quad (3.10-13)$$

Когда имеются только шумы, $P=0$ и

$$\begin{aligned} \overline{R} &= (2\psi_0)^{1/2} \Gamma\left(\frac{3}{2}\right) = \left(\frac{\psi_0 \pi}{2}\right)^{1/2} \\ \overline{R^2} &= 2\psi_0. \end{aligned} \quad (3.10-14)$$

Прежде чем идти дальше, в (3.10—11) удобно сделать следующие изменения обозначений:

$$v = \frac{R}{\psi_0^{1/2}}, \quad dv = \frac{dR}{\psi_0^{1/2}}, \quad a = \frac{P}{\psi_0^{1/2}}, \quad (3.10-15)$$

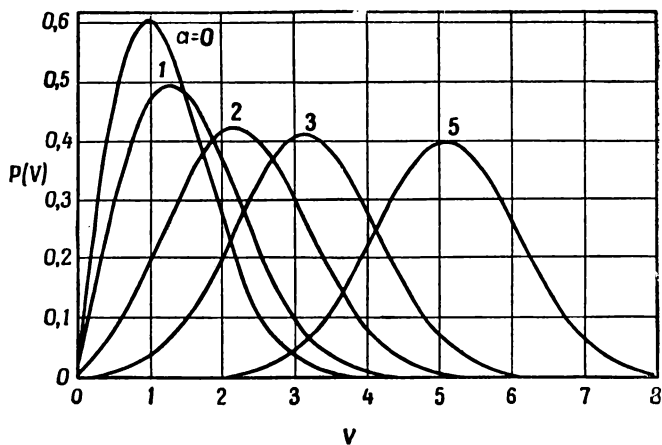
где a есть отношение амплитуды синусоидального тока к эффективному значению тока шумов.

Вместо случайной переменной R теперь имеется случайная переменная v , плотность вероятностей которой

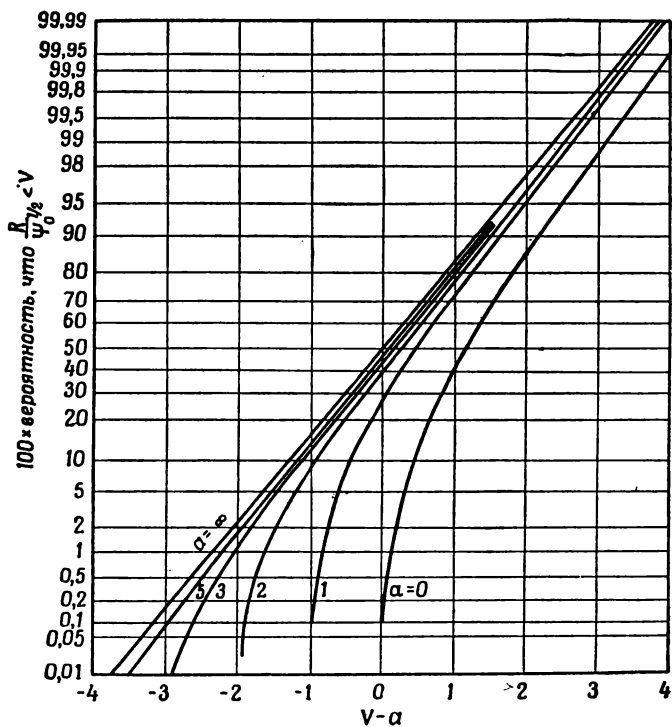
$$p(v) = v \exp\left(-\frac{v^2 + a^2}{2}\right) I_0(av). \quad (3.10-16)$$

Кривые $p(v)$ в функции от v представлены на фиг. 6 для значений $a=0, 1, 2, 3, 5$. Кривые, показывающие вероятность того, что v меньше, чем заданная величина, т. е. кривые распределения v , приведены на фиг. 7. Эти кривые были получены путем численного

¹⁾ Кривые этой функции приведены в «Таблицах функций» Янке и Эмде, 373, 1948, а некоторые ее свойства изложены в Приложении 4В.



Фиг. 6. Плотность вероятностей огибающей R тока $I(t) = P \cos pt + I_N$.



Фиг. 7. Функция распределения вероятностей огибающей R тока $I(t) = P \cos pt + I_N$.

интегрирования $p(v)$. Следующее полезное выражение для этой вероятности было дано У. Р. Беннетом в одной неопубликованной работе

$$\int_0^v p(u) du = \exp\left(-\frac{v^2 + a^2}{2}\right) \sum_{n=1}^{\infty} \left(\frac{v}{a}\right)^n I_n(av). \quad (3.10-17)$$

Это выражение получается путем интегрирования по частям, пользуясь соотношением

$$\int u^n I_{n-1}(au) du = \frac{u^n}{a} I_n(au).$$

При $av \gg 1$, но $1 \ll a-v$, Беннет показал, что (3.10-17) приводит к

$$\int_0^v p(u) du \approx \left(\frac{v}{2\pi a}\right)^{1/2} \frac{1}{a-v} \exp\left[-\frac{(v-a)^2}{2}\right] \times \\ \times \left[1 - \frac{3(a+v)^2 - 4v^2}{8av(a-v)^2} \dots\right] \quad (3.10-18)$$

Эта формула может быть также получена путем подстановки асимптотического разложения (3.10-19) для $p(v)$ в (3.10-17), выполнения дважды интегрирования по частям и пренебрежения членами высшего порядка.

• Когда av становится большим, $I_0(av)$ можно заменить его асимптотическим значением. Тогда выражение для $p(v)$ превращается в

$$p(v) \approx \left(1 + \frac{1}{8av}\right) \left(\frac{v}{2\pi a}\right)^{1/2} \exp\left[-\frac{(v-a)^2}{2}\right]. \quad (3.10-19)$$

Поэтому, когда a становится большим, либо v находится далеко у края кривой плотности вероятностей, распределение становится подобным нормальному закону. Нормальный закон определяет среднее значение P и стандартное отклонение $\phi_0^{1/2}$. Это стандартное отклонение таково же, что и стандартное отклонение мгновенных значений I_N .

Когда $av \gg 1$ и $a \gg |v-a|$, можно разложить коэффициент при показательной функции в (3.10-19) в ряд по $\frac{v-a}{a}$. Почленное интегрирование этого разложения дает, если пренебречь членами, по величине меньшими a^{-3} :

$$\int_0^v p(u) du \approx \frac{1}{2} + \frac{1}{2} \operatorname{erf} \frac{v-a}{\sqrt{2}} - \\ - \frac{1}{2a\sqrt{2\pi}} \left[1 - \frac{v-a}{4a} + \frac{1+(v-a)^2}{8a^2}\right] \exp\left[-\frac{(v-a)^2}{2}\right].$$

Когда I состоит из двух синусоидальных токов плюс шумы:

$$I = P \cos pt + Q \sin qt + I_N, \quad (3.10-20)$$

где угловые частоты p и q не имеют общего множителя, плотность вероятностей огибающей R равна

$$R \int_0^{\infty} r J_0(Rr) J_0(Pr) J_0(Qr) e^{-\frac{\phi_0 r^2}{2}} dr, \quad (3.10-21)$$

где $\phi_0 = \overline{I_N^2}$. При $Q=0$ интеграл может быть вычислен и тогда получается (3.10-11). Когда $P=Q=0$, получаем плотность вероятностей для R при наличии одних только шумов. Если вместо двух синусоидальных токов будут три, то в подинтегральное выражение следует поместить еще одну функцию Бесселя и т. д.

Для определения R удобно считать шумы сосредоточенными в сравнительно узком частотном диапазоне, а частоты синусоидальных токов — лежащими внутри этой полосы или вблизи нее. Как и в уравнениях (3.7-2) — (3.7-4), относим все члены к средней частоте диапазона $f_m = \omega_m / 2\pi$, пользуясь уравнениями типа

$$\begin{aligned} \cos pt &= \cos[(p - \omega_m)t + \omega_m t] = \\ &= \cos(p - \omega_m)t \cos \omega_m t - \sin(p - \omega_m)t \sin \omega_m t. \end{aligned}$$

Таким путем получим

$$V = A \cos \omega_m t - B \sin \omega_m t = R \cos(\omega_m t + \theta), \quad (3.10-22)$$

где A и B — сравнительно медленно изменяющиеся функции t , равные

$$\begin{aligned} A &= P \cos(p - \omega_m)t + Q \cos(q - \omega_m)t + \\ &\quad + \sum_n c_n \cos(\omega_n t - \omega_m t - \varphi_n), \\ B &= P \sin(p - \omega_m)t + Q \sin(q - \omega_m)t + \\ &\quad + \sum_n c_n \sin(\omega_n t - \omega_m t - \varphi_n) \end{aligned} \quad (3.10-23)$$

и

$$\begin{aligned} R^2 &= A^2 + B^2, \quad R > 0, \\ \operatorname{tg} \theta &= \frac{B}{A}. \end{aligned} \quad (3.10-24)$$

Как и следовало ожидать, уравнение (3.10-21) тесно связано с задачей о случайных смещениях и может быть получено из вывода Ключивера¹⁾, полагая, что шумы соответствуют весьма большому числу очень малых случайных смещений.

¹⁾ I. Н. Ватсон, «Теория бесселевых функций», ГИИЛ, 1949.

Другой путь вывода (3.10—21) основан на предположении, что $(p - \omega_m)t$, $(q - \omega_m)t$, $\varphi_1, \varphi_2, \dots$ суть независимые случайные углы. Характеристическая функция для A, B равна

$$\text{сред. } e^{iuA + ivB} = J_0(P\sqrt{u^2 + v^2}) J_0(Q\sqrt{u^2 + v^2}) e^{-\frac{\psi_0}{2}(u^2 + v^2)}.$$

Плотность вероятностей для A, B составит

$$\left(\frac{1}{2\pi}\right)^2 \int_{-\infty}^{+\infty} du \int_{-\infty}^{+\infty} dv e^{-iuA - ivB} \text{ сред. } e^{iuA + ivB}$$

Если произвести замену переменных

$$\begin{aligned} A &= R \cos \theta, & u &= r \cos \varphi, \\ B &= R \sin \theta, & v &= r \sin \varphi, \end{aligned}$$

то интегрирование по φ может быть выполнено. Двойной интеграл превращается в

$$\frac{1}{2\pi} \int_0^{\infty} r J_0(Pr) J_0(Qr) J_0(Rr) e^{-\frac{\psi_0}{2} r^2} dr.$$

Это ведет прямо к (3.10—21), если обратить внимание на то, что $dAdB = R dR d\theta$.

Кстати, если

$$I = Q(1 + k \cos pt) \cos qt + I_N,$$

где $p \ll q$, то подобные же соображения показывают, что плотность вероятностей для R равна

$$\frac{R}{2\pi} \int_0^{2\pi} d\alpha \int_0^{\infty} r J_0(Rr) J_0[Qr(1 + k \cos \alpha)] e^{-\frac{\psi_0}{2} r^2} dr,$$

где ω_m принято равным q . Интегрирование по r может быть выполнено. Это соотношение тесно связано с (3.10—11).

Возвращаясь теперь к случаю, когда I состоит из суммы двух синусоидальных токов плюс шумы, можно показать при помощи (3.10—21) и соотношения

$$\int_0^{\infty} R^{n+1} J_0(Rr) dR = \frac{2^{n+1} \Gamma\left(1 + \frac{n}{2}\right)}{r^{n+2} \Gamma\left(-\frac{n}{2}\right)},$$

что среднее значение R^n равно, если $-2 < \text{Re}(n) < -\frac{3}{2}$,

$$\overline{R^n} = \frac{2^{n+1} \Gamma\left(1 + \frac{n}{2}\right)}{\Gamma\left(-\frac{n}{2}\right)} \int_0^{\infty} r^{-n-1} J_0(Pr) J_0(Qr) e^{-\frac{\psi_0}{2} r^2} dr =$$

$$\begin{aligned}
&= (2\psi_0)^{n/2} \Gamma\left(\frac{n}{2} + 1\right) \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \frac{\left(-\frac{n}{2}\right)_{k+m} (-x)^k (-y)^m}{k! k! m! m!} = \\
&= (2\psi_0)^{n/2} \Gamma\left(\frac{n}{2} + 1\right) \sum_{k=0}^{\infty} \frac{\left(-\frac{n}{2}\right)_k (y-x)^k}{k! k!} P_k\left(\frac{x+y}{x-y}\right). \quad (3.10-25)
\end{aligned}$$

Представляется весьма вероятным, что этот результат может быть распространен путем аналитического продолжения на положительные целые значения n . В (3.10-25) применялись обозначения

$$\begin{aligned}
(\alpha)_0 &= 1, & (\alpha)_k &= \alpha(\alpha+1)\dots(\alpha+k-1), \\
x &= \frac{P^2}{2\psi_0}, & y &= \frac{Q^2}{2\psi_0}, \quad (3.10-26)
\end{aligned}$$

а полиномы Лежандра были обозначены через $P_k(z)$. Ряд сходится для всех значений P , Q и ψ_0 и обрывается, когда n есть четное положительное число.

Если x или y или оба вместе велики сравнительно с единицей, то можно из интеграла для $\overline{R^n}$ получить асимптотическое разложение, полагая $Q < P$, так что $y < x$:

$$\overline{R^n} \approx P^n \sum_{k=0}^{\infty} \frac{\left(-\frac{n}{2}\right)_k \left(-\frac{n}{2}\right)_k}{k! x^k} {}_2F_1\left(k - \frac{n}{2}, k - \frac{n}{2}; 1; \frac{y}{x}\right). \quad (3.10-27)$$

Когда n — четное положительное число, этот ряд обрывается и дает такое же выражение, что и (3.10-25). Когда n — нечетное целое число, функция ${}_2F_1$ может быть выражена через полные эллиптические функции E и K модуля $y^{1/2}, x^{-1/2}$.

$$\begin{aligned}
{}_2F_1\left(-\frac{1}{2}, -\frac{1}{2}; 1; \frac{y}{x}\right) &= \frac{4}{\pi} E - \frac{2}{\pi} \left(1 - \frac{y}{x}\right) K, \\
{}_2F_1\left(\frac{1}{2}, \frac{1}{2}; 1; \frac{y}{x}\right) &= \frac{2}{\pi} K. \quad (3.10-28)
\end{aligned}$$

Высшие члены могут быть вычислены из

$$\begin{aligned}
a(1-z)^2 {}_2F_1(a+1, a+1; 1; z) &= \\
&= (2a-1)(1+z) {}_2F_1(a, a; 1; z) + \\
&+ (1-a) {}_2F_1(a-1, a-1; 1; z), \quad (3.10-29)
\end{aligned}$$

что является частным случаем соотношения

$$\begin{aligned} ab(\gamma+1)(1-z)^2 {}_2F_1(a+1, b+1; c; z) = \\ = A {}_2F_1(a, b; c; z) - \\ - (\gamma-1)(c-a)(c-b) {}_2F_1(a-1, b-1; c; z), \end{aligned} \quad (3.10-30)$$

где $\gamma=c-a-b$, и

$$A = (i^2 - 1)\gamma + (1-z)[(\gamma-1)(c-b)(b-1) + (\gamma+1)a(c-a-1)].$$

Хотя из данного выражения это и не видно, но A действительно симметрично по a и b . Симметричная форма может быть получена при использовании выражения, которое находится, если в (3.10—30) положить $z=0$.

3.11. НЕКОТОРЫЕ ДОПОЛНЕНИЯ О ТОКАХ ДРОБОВОГО ЭФФЕКТА

В большинстве выводов этой главы в качестве исходных использовались выражения (2.8—1) или (2.8—6). Покажем сейчас, что представление токов дробового эффекта, применявшееся в главе I, может быть также взято как исходное.

Например, допустим, что надо найти двумерное распределение $I(t)$ и $I(t+\tau)$, рассмотренное в разделе 3.2. Это — частный случай распределения двух переменных

$$\begin{aligned} I(t) &= \sum_{k=-\infty}^{+\infty} F(t-t_k), \\ J(t) &= \sum_{k=-\infty}^{+\infty} G(t-t_k), \end{aligned} \quad (3.11-1)$$

где теперь предположим

$$\int_{-\infty}^{+\infty} F(t) dt = \int_{-\infty}^{+\infty} G(t) dt = 0, \quad (3.11-2)$$

чтобы средние значения I и J могли быть равны нулю. Чтобы получить $I(t+\tau)$ из $J(t)$, положим $G(t)=F(t+\tau)$.

Распределение I и J может быть найдено во многом подобно тому, как в разделе 1.4 было найдено распределение для одного I . Характеристическая функция распределения равна

$$f(u, v) = \text{сред. } e^{iuI+ivJ} = \exp \nu \int_{-\infty}^{+\infty} \left\{ \exp [iuF(t)+ivG(t)] - 1 \right\} dt, \quad (3.11-3)$$

где ν есть ожидаемое число событий (т. е. для дробового эффекта — попаданий на анод электронов) в 1 сек. Плотность вероятностей для I и J равна

$$\frac{1}{4\pi^2} \int_{-\infty}^{+\infty} du \int_{-\infty}^{+\infty} dv e^{-iuI - ivJ} f(u, v). \quad (3.11-4)$$

Семиинварианты $\lambda_{m,n}$ находятся из

$$\log f(u, v) = \sum_{m, n=1}^k \frac{\lambda_{m,n}}{m!n!} (iu)^m (iv)^n + \eta [(iu)^k, (iv)^k]$$

и равны

$$\lambda_{m,n} = \nu \int_{-\infty}^{+\infty} F^m(t) G^n(t) dt. \quad (3.11-5)$$

Когда $\nu \rightarrow \infty$, распределение I и J сходится к двумерному нормальному закону. Приближение к этому нормальному закону можно найти, следуя методике раздела 1.6. Из допущения (3.11-2) следует, что λ_{10} и λ_{01} равны нулю. Из соотношений между моментами второго порядка и семиинвариантами λ имеем

$$\mu_{11} = \lambda_{20} + \lambda_{10}^2 = \nu \int_{-\infty}^{+\infty} F^2(t) dt,$$

$$\mu_{12} = \lambda_{11} + \lambda_{10} \lambda_{01} = \nu \int_{-\infty}^{+\infty} F(t) G(t) dt, \quad (3.11-6)$$

$$\mu_{22} = \lambda_{02} + \lambda_{01}^2 = \nu \int_{-\infty}^{+\infty} G^2(t) dt.$$

Здесь индексы при μ отличаются от индексов при λ , причем это изменение сделано для того, чтобы привести индексы в соответствие с разделами 2.9 и 2.10, так что можно сразу написать нормальное распределение.

Глава IV

ПРОХОЖДЕНИЕ СИГНАЛА И ШУМОВ ЧЕРЕЗ НЕЛИНЕЙНЫЕ УСТРОЙСТВА

Здесь будут рассмотрены две проблемы, относящиеся к прохождению шумов через детекторы или другие нелинейные устройства. Первая относится к статистическим свойствам тока на выходе нелинейного устройства, т. е. среднему значению тока, флуктуациям вокруг этого среднего значения и т. д.

Вторая проблема может быть сформулирована более определенно: дано нелинейное устройство, на вход которого подается либо одно только напряжение шумов, либо напряжения шумов и сигнала. Каков будет энергетический спектр на выходе?

При написании этой главы полезным оказалось ознакомление со статьей Беннета ¹⁾, а также с рукописью вскоре выходящей статьи Миддльтона ²⁾.

4.1. ТОК НИЗКОЙ ЧАСТОТЫ НА ВЫХОДЕ КВАДРАТИЧНОГО ДЕТЕКТОРА

Пусть ток на выходе устройства I следующим образом связан с входным напряжением V :

$$I = \alpha V^2, \quad (4.1-1)$$

где $\alpha = \text{const}$. Если энергетический спектр V ограничен сравнительно узким диапазоном частот, то энергетический спектр I состоит из двух частей. Одна часть сосредоточена вокруг удвоенной средней частоты спектра V , а другая — вокруг нулевой частоты. Нас интересует низкочастотная часть. Ток, соответствующий этой части спектра, будет обозначаться I_{II} . Это ток, который должен протекать на выходе, если включить фильтр нижних частот для отфильтрования верхней части спектра. Удобно разделить I_{II} на две составляющие

$$I_{II} = I_{dc} + I_{lf}, \quad (4.1-2)$$

¹⁾ Цитирована ранее (раздел 3.10). В следующих разделах при ссылках на статью Беннета и рукопись Миддльтона упоминаются только фамилии авторов.

²⁾ Автор, повидимому, имеет в виду работу Миддльтона «Реакция линейного и квадратичного детекторов на флуктуационные шумы», опубликованную в *Journ. Appl. Phys.*, 17, 778, октябрь 1946. (Прим. ред.)

где индексы обозначают, соответственно, «полная низкая» частота (I_f), «постоянный ток» (I_{dc}) и «низкая частота» (I_l). Имеем

$$I_{dc} = \overline{I_{il}},$$

$$\overline{I_{if}^2} = \overline{(I_{il} - I_{dc})^2} = \overline{I_{il}^2} - I_{dc}^2. \quad (4.1-3)$$

Простейший способ нахождения I_{dc} состоит в возведении в квадрат данного выражения для V и выделении членов, не зависящих от времени. Поэтому, если

$$V = P \cos pt + Q \cos qt + \dot{V}_N, \quad (4.1-4)$$

то получим

$$I_{dc} = \alpha \left(\frac{P^2}{2} + \frac{Q^2}{2} + \overline{V_N^2} \right). \quad (4.1-5)$$

I_{if} может быть, конечно, определен путем выделения низкочастотных членов. Вместо этого здесь будет развит на примере квадратичного детектора, а в следующем разделе — линейного детектора общий метод исследования статистических свойств тока на выходе нелинейного устройства, когда входное напряжение ограничено относительно узким диапазоном частот.

Если низкочастотные составляющие спектра целиком пропускаются фильтрами, то

$$I_{il} = \alpha \frac{R^2}{2}, \quad (4.1-6)$$

где R — огибающая напряжения V . Плотность вероятностей и статистические свойства I_{il} могут быть найдены из этого уравнения, если известна функция распределения R^2 ¹⁾. Прежде чем рассматривать эти свойства, докажем справедливость (4.1-6).

Уравнение (4.1-6) является частным случаем более общего решения, полученного в разделе 4.3. Его справедливость может быть показана рассмотрением примера

$$V = P \cos pt + Q \cos qt + V_N, \quad (4.1-4)$$

где $f_p = p/2\pi$ и $f_q = q/2\pi$ лежат внутри, либо примыкают к частотному диапазону напряжения шумов V_N .

¹⁾ Если часть низкочастотного спектра не пропускается, то вопрос усложняется. I_{dc} может быть найден, как и выше, но для нахождения $\overline{I_{if}^2}$ необходимо сначала определить энергетический спектр I (раздел 4.5) и затем проинтегрировать по соответствующей его части. Что касается распределения I_{if} , то пока можно только утверждать, что оно находится между распределением по уравнению (4.1-6) и нормальным распределением, которого оно достигает, когда только узкий участок низкочастотного спектра пропускается фильтром звуковых частот (раздел 4.3).

Применяя формулы типа

$$\begin{aligned} \cos pt = \cos [(p - \omega_m)t + \omega_m t] = \cos (p - \omega_m)t \cos \omega_m t - \\ - \sin (p - \omega_m)t \sin \omega_m t, \end{aligned} \quad (4.1-7)$$

можно отнести все члены к средней частоте диапазона $f_m = \omega_m / 2\pi$, как это сделано в (3.7-2) и (3.7-4). Таким образом, получим

$$V = A \cos \omega_m t - B \sin \omega_m t = R \cos (\omega_m t + \theta), \quad (4.1-8)$$

где A и B — относительно медленно изменяющиеся функции t , равные

$$A = P \cos (p - \omega_m)t + Q \cos (q - \omega_m)t + \sum_n c_n \cos (\omega_n t - \omega_m t - \varphi_n),$$

$$B = P \sin (p - \omega_m)t + Q \sin (q - \omega_m)t + \sum_n c_n \sin (\omega_n t - \omega_m t - \varphi_n),$$

а

$$\begin{aligned} R^2 = A^2 + B^2, \quad R > 0, \\ \operatorname{tg} \theta = B/A. \end{aligned} \quad (4.1-9)$$

Такое определение R было также дано в (3.10-22, 23, 24). Огибающая V есть R , а выходной ток равен

$$I = \alpha R^2 \left[\frac{1}{2} + \frac{1}{2} \cos (2\omega_m t + 2\theta) \right]. \quad (4.1-10)$$

Так как R — медленно изменяющаяся функция времени, то такой же функцией является и R^2 . Энергетический спектр R^2 ограничен частотами, значительно более низкими, чем $2f_m$, и в результате энергетический спектр $R^2 \cos (2\omega_m t + 2\theta)$ сосредоточен вокруг $2f_m$. Поэтому единственным членом в уравнении для I , соответствующим низкочастотному выходному току, является $\alpha R^2 / 2$, что и нужно было доказать.

Возвратимся теперь к статистическим свойствам I_{it} . Во-первых, рассмотрим случай, когда напряжение V состоит только из одних шумов $V = V_N$, так что плотность вероятностей для огибающей R равна

$$\frac{R}{\psi_0} e^{-R^2/2\psi_0}, \quad (3.7-10)$$

где

$$\psi_0 = [\text{эфф. знач. } V_N]^2 = \overline{V_N^2}, \quad (4.1-11)$$

откуда

$$I_{dc} = \overline{I_{it}} = \frac{\alpha \overline{R^2}}{2} = \int_0^\infty \frac{\alpha R^2}{2} \frac{R}{\psi_0} e^{-R^2/2\psi_0} dR = \alpha \psi_0.$$

$$\overline{I_{if}^2} = \overline{I_{it}^2} - I_{dc}^2 = \int_0^\infty \frac{\alpha^2 R^5}{4 \psi_0} e^{-R^2/2\psi_0} dR - I_{dc}^2 = \alpha^2 \psi_0^2. \quad (4.1-12)$$

Во-вторых, рассмотрим случай, когда

$$V = V_N + P \cos pt, \quad (4.1-13)$$

где $\rho/2\pi$ лежит вблизи полосы частот шумов V_N . Плотность вероятностей для огибающей R равна

$$\frac{R}{\psi_0} \exp \left[-\frac{R^2 + P^2}{2\psi_0} \right] I_0 \left(\frac{RP}{\psi_0} \right). \quad (3.10-11)$$

Отсюда и из уравнений (3.10-12), (3.10-13) найдем

$$I_{dc} = \frac{\alpha \bar{R}^2}{2} = \alpha \psi_0 + \frac{\alpha P^2}{2}, \quad (4.1-14)$$

$$\bar{I}_{ii}^2 = \frac{\alpha^2}{4} \bar{R}^4 = \alpha^2 \left[2\psi_0^2 + 2P^2\psi_0 + \frac{P^4}{4} \right],$$

$$\bar{I}_{if}^2 = \bar{I}_{ii}^2 - I_{dc}^2 = \alpha^2 (\psi_0 + P^2) \psi_0. \quad (4.1-15)$$

В (4.1-14) ψ_0 есть средний квадрат напряжения шумов V_N , а $P^2/2$ — средний квадрат сигнала. Эти уравнения показывают, что I_{dc} и эффективное значение I_{if} не зависят от распределения энергетического спектра шумов V_N до тех пор, пока входное напряжение V ограничено относительно узким диапазоном частот. Другими словами, хотя это распределение и влияет на выходной энергетический спектр, но оно не влияет на постоянную составляющую и эффективное значение I_{if} , когда ψ_0 и P заданы. То, что это справедливо для большой группы нелинейных устройств, было указано Миддлтоном (см. конец раздела 4.9).

Когда входное напряжение равно

$$V = V_N + P \cos pt + Q \cos qt, \quad (4.1-4)$$

$p \neq q$, получим из уравнения (3.10-25)

$$I_{dc} = \frac{\alpha}{2} \bar{R}^2 = \alpha \left(\psi_0 + \frac{P^2}{2} + \frac{Q^2}{2} \right),$$

$$\bar{I}_{ii}^2 = \frac{\alpha^2}{4} \bar{R}^4, \quad (4.1-16)$$

$$\bar{I}_{if}^2 = \alpha^2 \left[\psi_0^2 + P^2\psi_0 + Q^2\psi_0 + \frac{P^2Q^2}{2} \right].$$

4.2. ТОК НИЗКОЙ ЧАСТОТЫ НА ВЫХОДЕ ЛИНЕЙНОГО ДЕТЕКТОРА

В случае линейного детектора

$$I = \begin{cases} 0, & V < 0, \\ \alpha V, & V > 0, \end{cases} \quad (4.2-1)$$

и выходной ток низкой частоты, считая, что фильтр звуковой частоты отсутствует, равен ¹⁾

$$I_{ii} = \frac{\alpha R}{\pi} \quad (4.2-2)$$

¹⁾ Автор рассматривает случай безинерционного детектирования. (Прим. ред.)

Здесь, как и в аналогичном случае (4.1—6) при квадратичном детекторе, предполагается, что подводимый сигнал и шумы лежат внутри относительно узкого диапазона частот. Эта формула может быть использована для вычисления плотности вероятностей и статистических свойств I_{it} , если известны необходимые данные об огибающей R приложенного напряжения.

Справедливость (4.2—2) может быть доказана рассмотрением выходного тока I . Он состоит из положительных полуволн αV , огибающая I тождественна огибающей αV . Однако площадь, лежащая под кривой I , составляет только $1/\pi$ площади, лежащей под кривой αR ; именно таково соотношение площади, ограниченной кривой $\sin x$, к площади прямоугольника, имеющего единичную высоту и длину 2π . В результате ток низкой частоты изменяется по закону $\alpha R/\pi$.

Если V состоит из синусоидальных колебаний и шумов

$$V = V_N + P \cos pt, \quad (4.1-13)$$

то среднее значение I_{it} равно

$$\begin{aligned} I_{dc} &= \frac{\alpha}{\pi} \bar{R} = \alpha \left(\frac{\psi_0}{2\pi} \right)^{1/2} {}_1F_1 \left(-\frac{1}{2}; 1; -\frac{P^2}{2\psi_0} \right) = \\ &= \alpha \left(\frac{\psi_0}{2\pi} \right)^{1/2} e^{-x/2} \left[(1+x) I_0 \left(\frac{x}{2} \right) + x I_1 \left(\frac{x}{2} \right) \right] \end{aligned} \quad (4.2-3)$$

где I_0, I_1 — бесселевы функции мнимого аргумента,

$$x = \frac{P^2}{2\psi_0} = \frac{\text{сред. мощность синусоид. колебаний}}{\text{сред. мощность шумов}}, \quad (4.2-4)$$

а ψ_0 есть среднее значение V_N^2 . Уравнение (4.2—3) следует из формул (3.10—12) и (4.В—9)¹⁾. Если x велико, то асимптотическое разложение (4.В—3) для ${}_1F_1$ дает

$$I_{dc} \approx \frac{\alpha^2}{\pi^2} \left[P + \frac{\psi_0}{2P} + \frac{\psi_0^2}{8P^3} + \dots \right] \quad (4.2-5)$$

Подобным же образом средний квадрат тока I_{it} равен

$$\overline{I_{it}^2} = \frac{\alpha^2}{\pi^2} \overline{R^2} = \frac{\alpha^2}{\pi^2} (P^2 + 2\psi_0), \quad (4.2-6)$$

а средний квадрат тока низкой частоты I_{lf} (без постоянной составляющей) равен

$$\overline{I_{lf}^2} = \overline{I_{it}^2} - I_{dc}^2.$$

¹⁾ См. Приложение 4В. (Прим. ред.)

Если x велико, то

$$\bar{I}_{lf}^2 \approx \frac{\alpha^2}{\pi^2} \left(\psi_0 - \frac{\psi_0^2}{2P^2} \right) = \frac{\alpha^2}{\pi^2} \psi_0 \left(1 - \frac{1}{4x} \right), \quad (4.2-7)$$

а если $x=0$,

$$\bar{I}_{lf}^2 = \frac{\alpha^2}{\pi^2} \psi_0 \left(2 - \frac{\pi}{2} \right). \quad (4.2-8)$$

Кривые I_{dc} приведены на фиг. 1, 2 и 3 статьи Беннета. Он также приводит на фиг. 4 кривые, показывающие зависимость \bar{I}_{lf}^2 от x . Из них следует, что влияние комбинационных членов высшего порядка мало, когда I_{lf} находится путем сложения низкочастотных комбинационных тонов.

Если напряжение V состоит из двух синусоидальных колебаний и шумов

$$V = V_N + P \cos pt + Q \cos qt, \quad (4.1-4)$$

то в среднее значение I_{ll} по уравнению (3.10-25) входит двойная функция ${}_1F_1$

$$\begin{aligned} I_{dc} &= \frac{\alpha}{\pi} \bar{R} = \alpha \left(\frac{\psi_0}{2\pi} \right)^{1/2} \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-\frac{1}{2})_{k+m}}{k! k! m! m!} (-x)^k (-y)^m = \\ &= \alpha \left(\frac{\psi_0}{2\pi} \right)^{1/2} \sum_{k=0}^{\infty} \frac{(-\frac{1}{2})_k}{k! k!} (y-x)^k P_k \left(\frac{x+y}{x-y} \right), \end{aligned} \quad (4.2-9)$$

где

$$x = \frac{P^2}{2\psi_0}; \quad y = \frac{Q^2}{2\psi_0}, \quad P_k(z) \text{ — полином Лежандра.} \quad (4.2-10)$$

Если x велико и $y < x$, то из (3.10-27) получим асимптотическое выражение

$$I_{dc} \approx \frac{\alpha}{\pi} P \sum_{k=0}^{\infty} \frac{(-\frac{1}{2})_k (-\frac{1}{2})_k}{k! x^k} {}_2F_1 \left(k - \frac{1}{2}, k - \frac{1}{2}; 1; \frac{y}{x} \right). \quad (4.2-11)$$

Функция ${}_2F_1$ может быть выражена через полные эллиптические функции E и K модуля $y^{1/2} x^{-1/2}$. Поэтому

$$\begin{aligned} {}_2F_1 \left(-\frac{1}{2}, -\frac{1}{2}; 1; \frac{y}{x} \right) &= \frac{4}{\pi} E - \frac{2}{\pi} \left(1 - \frac{y}{x} \right) K \\ {}_2F_1 \left(\frac{1}{2}, \frac{1}{2}; 1; \frac{y}{x} \right) &= \frac{2}{\pi} K, \end{aligned} \quad (3.10-28)$$

а высшие члены могут быть вычислены по рекуррентному уравнению (3.10-29). Первый член $k=0$ в (4.2-11) дает I_{dc} , когда шумы отсутствуют.

Средний квадрат тока I_{II} равен

$$\overline{I_{II}^2} \approx \frac{\alpha^2}{\pi^2} \overline{R^2} = \frac{\alpha^2}{\pi^2} (2\psi_0 + P^2 + Q^2). \quad (4.2-12)$$

Из этого уравнения, а также из выражения для I_{dc} может быть найдено эффективное значение тока низкой частоты (за вычетом постоянной составляющей). Например, когда шумы невелики,

$$\begin{aligned} \overline{I_{II}^2} \approx \frac{\alpha^2}{\pi^2} \left\{ P^2 + Q^2 - \left[P {}_2F_1 \left(-\frac{1}{2}, -\frac{1}{2}; 1; \frac{y}{x} \right) \right]^2 + \right. \\ \left. + 2\psi_0 \left[1 - {}_2F_1 \left(-\frac{1}{2}, -\frac{1}{2}; 1; \frac{y}{x} \right) \frac{K}{\pi} \right] \right\}, \quad (4.2-13) \end{aligned}$$

Член, не зависящий от ψ_0 , дает средний квадрат тока низкой частоты в отсутствие шумов. Когда Q уменьшается до нуля, (4.2-13) сводится к главному члену (4.2-7), как и должно быть. Если $P=Q$, то формула негодна и необходимо пользоваться асимптотическим значением

$$I_{dc} = \alpha \left(\frac{\psi_0}{2\pi} \right)^{1/4} \sum_{k=0}^{\infty} \frac{(-\frac{1}{2})_k (2k)!}{(k!)^4} (-x)^k.$$

Ввиду нестроого характера выводов в разделе 3.10 представляется ценной численная проверка эквивалентности уравнений (4.2-9) и (4.2-11). Для этого во второй ряд (4.2-9) подставляем значения $x=4$, $y=3$. Оказывается, что наибольший член в суммировании имеет место при $k=11$. В общем, учитываем 24 члена. В результате получаем

$$\frac{\overline{R}}{\sqrt{2\psi_0}} = 2,5502.$$

Для тех же значений x и y асимптотический ряд (4.2-11) дает

$$2,40 + 0,171 + 0,075 + 0,052 + \dots$$

Если остановиться перед третьим членом, то сумма равна 2,57. При включении наименьшего члена получим 2,65. Соответствие результатов показывает, что (4.2-11) действительно является асимптотическим разложением (4.2-9).

Если входное напряжение имеет вид

$$V = Q(1 + k \cos pt) \cos qt + V_N,$$

то можно воспользоваться уравнением

$$\overline{R^n} = (2\psi_0)^{n/2} \Gamma\left(1 + \frac{n}{2}\right) \frac{1}{2\pi} \int_0^{2\pi} {}_1F_1 \left[-\frac{n}{2}; 1; -y(1 + k \cos \theta)^2 \right] d\theta,$$

где R — огибающая по отношению к частоте $q/2\pi$, а y находится из уравнения (4.2—10). Интеграл может быть вычислен путем представления ${}_1F_1$ в виде степенного ряда и почленного интегрирования, пользуясь соотношением

$$\frac{1}{2\pi} \int_0^{2\pi} (1 + k \cos \theta)^l \cos m\theta \, d\theta = \frac{(-l)_m}{2^m m!} (-k)^m {}_2F_1 \left[\frac{m-l}{2}, \frac{m-l+1}{2}; m+1; k^2 \right], \quad (4.2-15)$$

где m — неотрицательное целое число, l — любое число,

$$(\alpha)_m = \alpha(\alpha+1)\cdots(\alpha+m-1), \quad (\alpha)_0 = 1, \quad (0)_0 = 1.$$

Интеграл может быть также взят при помощи присоединенных функций Лежандра.

Применяя методы раздела 3.10 к (4.2—14), приходим к следующим выражениям:

$$\begin{aligned} \bar{R}^2 &= Q^2 \left(1 + \frac{k^2}{2} \right) + 2\psi_0, \\ \bar{R} &\approx Q \sum_{s=0}^{\infty} \frac{(-\frac{1}{2})_s (-\frac{1}{2})_s}{s! y^s} {}_2F_1 \left(s - \frac{1}{2}, s; 1; k^2 \right) \end{aligned} \quad (4.2-16)$$

где асимптотический ряд применим, если y очень велико, а k не слишком близко к единице. Из этих выражений

$$\bar{I}_{if}^2 \approx \frac{\alpha^2}{\pi^2} \left\{ Q^2 \frac{k^2}{2} + \psi_0 [2 - (1 - k^2)^{-1/2}] + \dots \right\}. \quad (4.2-17)$$

Сопоставление коэффициентов при ψ_0 в (4.2—17) со сплошным участком выходного энергетического спектра не будет правильным. Основная составляющая, вносимая в \bar{I}_{if}^2 сплошным участком энергетического спектра, есть $\alpha^2 \psi_0 / \pi^2$, т. е. как и в (4.2—7), если $k=0$. Разница между этим значением и соответствующим членом в (4.2—17) появляется, повидимому, из-за того, что амплитуда выделенного сигнала не точно равна $\alpha Q k / \pi$, но оказывается измененной вследствие присутствия шумов. Это обстоятельство можно было ожидать по физическим соображениям, так как изменение в уравнении (4.2—7) P , скажем удвоение, незначительно сказывается на токе \bar{I}_{if}^2 в (4.2—7), который обусловлен исключительно сплошным участком шумового спектра. Модулированная волна может рассматриваться как результат медленных изменений P .

4.3. НЕКОТОРЫЕ СТАТИСТИЧЕСКИЕ СВОЙСТВА ТОКА НА ВЫХОДЕ НЕЛИНЕЙНОГО УСТРОЙСТВА ОБЩЕГО ТИПА

Рассматриваемая задача такова: дано нелинейное устройство, ток на выходе которого связан с входным напряжением соотношением

$$I = \frac{1}{2\pi} \int_C F(iu) e^{iVu} du, \quad (4.A-1)$$

которое приведено в Приложении 4A¹). Пусть входное напряжение V наряду с сигналом содержит и шумы. Выберем для исследования на выходе некоторую полосу частот. Каковы будут статистические свойства тока в этой полосе частот?

Решение этой задачи в таком общем виде представляется трудным. Однако два следующих утверждения будут справедливы.

1. По мере сужения полосы частот на выходе статистические свойства соответствующего тока приближаются к свойствам шумовых токов, рассмотренным в главе III (полагая, что гармоники сигнала не попадают в эту полосу частот). В частности, мгновенные значения токов распределены по нормальному закону.

2. Если входное напряжение V ограничено относительно узкой полосой частот, то энергетический спектр выходного тока I сосредоточен вокруг нулевой (постоянная составляющая), 1-й, 2-й и т. д. гармоник средней частоты полосы частот напряжения V . Выходной ток низкой частоты, включая и постоянную составляющую, равен

$$I_n = A_0(R) = \frac{1}{2\pi} \int_C F(iu) J_0(uR) du, \quad (4.3-11)$$

где R — огибающая V

Огибающая n -й гармоники выходного тока при $n > 0$ равна

$$A_n(R) = \frac{1}{\pi} \int_C F(iu) J_n(uR) du. \quad (4.3-1)$$

Тогда

$$I = \sum_{n=0}^{\infty} A_n(R) \cos(n\omega_m t + n\theta), \quad (4.3-9)$$

где $f_m = \omega_m/2\pi$ — средняя частота полосы частот напряжения V , а θ — сравнительно медленно изменяющийся фазовый угол. Результаты, полученные в разделах 4.1 и 4.2, являются частными случаями этого уравнения.

Заклучение о том, что мощность шумов в каждой полосе частот на выходе (в полосе частот, соответствующей данной гармонике f_m) зависит только от $V_N^2 = \psi_0$ и не зависит от спектра V_N , где V_N — составляющая напряжения шумов во входном напряжении V , может быть также получено из (4.3—9). Замечаем, что полная мощность в n -й полосе частот зависит только от среднего квадрата

¹) Здесь $F(iu)$ представляет собой преобразование Фурье вольтамперной характеристики нелинейного устройства. (Прим. ред.)

оггибающей $A_n(R)$ и что в плотность вероятностей оггибающей R на-
пряжения шумов V_N входит только через ψ_0 .

Обоснование первого утверждения не является вполне удов-
летворительным. Оно сводится к следующему. Выходной ток I
может быть разделен на две составляющие. Одна, обусловленная
сигналом, состоит из синусоидальных членов. Другая является
шумовой составляющей.

Здесь будет рассматриваться последняя; обозначим ее I_N .
Корреляция между двумя значениями I_N , разделенными проме-
жутком времени, стремится к нулю, когда этот интервал стано-
вится большим. Пусть τ будет промежуток времени, достаточно
большой для обеспечения существенной независимости двух зна-
чений I_N . Выберем промежуток времени T достаточно большим
так, чтобы он заключал в себе много интервалов продолжитель-
ностью τ . Разложив I_N в этом промежутке в ряд Фурье, получим

$$I_N = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos \frac{2\pi n t}{T} + b_n \sin \frac{2\pi n t}{T} \right) \quad (4.3-2)$$

$$a_n - i b_n = \frac{2}{T} \int_0^T e^{-i 2\pi n t / T} I_N(t) dt.$$

Пусть рассматриваемая полоса частот простирается от $f_0 - \beta/2$
до $f_0 + \beta/2$ и пусть

$$T \left(f_0 - \frac{\beta}{2} \right) = n_1, \quad T \left(f_0 + \frac{\beta}{2} \right) = n_2, \quad (4.3-3)$$

где n_1 и n_2 — целые числа. Число составляющих тока в этой полосе
частот будет $(n_2 - n_1)$. Полагаем β существенно меньшим, чем $1/\tau$.
Выходной ток в этой полосе частот

$$I'_N = \sum_{n=n_1}^{n_2} \left(a_n \cos \frac{2\pi n t}{T} + b_n \sin \frac{2\pi n t}{T} \right), \quad (4.3-4)$$

где

$$a_n - i b_n = \frac{2}{T} \int_0^T e^{-i 2\pi \left(\frac{n}{T} - f_0 \right) t} e^{-i 2\pi f_0 t} I_N(t) dt, \quad (4.3-5)$$

$$n = \frac{n_1 + n_2}{2} + n - \frac{n_1 + n_2}{2} = f_0 T + (n - f_0 T).$$

Полоса частот выбрана настолько узкой, что

$$n_2 - n_1 \ll \frac{T}{\tau} \quad \text{или} \quad \beta \tau \ll 1. \quad (4.3-6)$$

Это дает возможность написать приближенно

$$a_n - ib_n = \sum_{r=1}^{r_1} e^{-i2\pi\left(\frac{n}{T} - f_0\right)r\tau} \frac{2}{T} \int_{(r-1)\tau}^{r\tau} e^{-i2\pi f_0 t} I_N(t) dt,$$

где $r_1 = \frac{T}{\tau}$, а T выбирается таким, чтобы r_1 было целым числом. Допустим, что так проделано с большим числом промежутков длительностью каждый T . Тогда $I_N(t)$ будет различным при переходе от интервала к интервалу. Группа интегралов для $r=1$ дает ряд значений, которые можно рассматривать как определяющие распределение комплексной случайной переменной, скажем x_1 . Аналогично ряд интегралов для $r=2$ определяет распределение второй случайной переменной x_2 и т. д. вплоть до x_{r_1} . Вследствие выбора τ настолько большим, что значение $I_N(t)$ в данном интеграле практически не зависит от его значений в других интегралах, можно сказать, что x_1, x_2, \dots, x_{r_1} являются независимыми.

Получаем

$$a_{n_1} - ib_{n_1} = \sum_{r=1}^{r_1} e^{-i2\pi\left(\frac{n}{T} - f_0\right)r\tau} x_r,$$

$$a_{n_1+1} - ib_{n_1+1} = \sum_{r=1}^{r_1} e^{-i2\pi\left(\frac{n_1+1}{T} - f_0\right)r\tau} x_r,$$

$$a_{n_2} - ib_{n_2} = \sum_{r=1}^{r_1} e^{-i2\pi\left(\frac{n_2}{T} - f_0\right)r\tau} x_r.$$

Если $n_2 - n_1 \ll r_1$, как это предполагалось в (4.3—6), то можно применить центральную предельную теорему для доказательства того, что $a_{n_1}, b_{n_1}, a_{n_1+1}, \dots, a_{n_2}, b_{n_2}$ стремятся стать независимыми и нормально распределенными вокруг нуля, по мере того, как ширина полосы $\beta \rightarrow 0$, а $T \rightarrow \infty$ (и, следовательно, $r_1 \rightarrow \infty$) таким образом, что $(n_2 - n_1)$ поддерживается постоянным. При этом используется то обстоятельство, что ток $I_N(t)$ таков, что вещественные и мнимые части x_1, x_2, \dots, x_r все имеют то же самое среднее значение и стандартное отклонение. Удобно полагать $f_0 T$ целым числом.

Поэтому, когда ширина полосы частот β достигает нуля, выходной ток в этой полосе I'_N , определяемый (4.3—4), может быть представлен таким же образом, как (2.8—1), т. е. как это было сделано при изучении шумового тока в главе III. Следовательно, статистические свойства I'_N должны сходиться к свойствам рассмотренного там шумового тока. Например, распределение вероятностей I'_N сходится к нормальному закону.

В нашем рассмотрении было положено, что $\beta\tau \ll 1$. Если напряжение V , подводимое к нелинейному устройству, ограничено относительно узким частотным диапазоном, скажем $f_b - f_a$, то промежуток времени τ (выбранный выше так, что $I(t)$ и $I(t+\tau)$ существенно независимы) может быть выбран порядка $1/(f_b - f_a)$. В этом случае I'_N ведет себя подобно шумовому току, если $\beta/(f_b - f_a)$ значительно меньше единицы.

Теперь обратимся ко второму утверждению, сделанному в начале этого раздела. Пусть приложенное напряжение ограничено сравнительно узким диапазоном частот так, чтобы оно могло быть представлено уравнением (4.1—8) раздела 4.1,

$$V = R \cos(\omega_m t + \theta), \quad R \geq 0, \quad (4.1-8)$$

где $f_m = \frac{\omega_m}{2\pi}$ есть некоторая опорная частота внутри полосы частот, а R и θ —функции времени, медленно изменяющиеся по сравнению с $\cos \omega_m t$. Через R обозначена огибающая V

Из уравнения (4А—1)

$$I = \frac{1}{2\pi} \int_C F(iu) e^{iuR \cos(\omega_m t + \theta)} du. \quad (4.3-7)$$

Разложим подинтегральное выражение при помощи соотношения

$$e^{ix \cos \varphi} = \sum_{n=0}^{\infty} \epsilon_n i^n \cos n\varphi J_n(x), \quad (4.3-8)$$

где ϵ_0 есть 1, ϵ_n есть 2, когда $n > 0$, а $J_n(x)$ —функция Бесселя. Поэтому

$$I = \sum_{n=0}^{\infty} A_n(R) \cos(n\omega_m t + n\theta), \quad (4.3-9)$$

где

$$A_n(R) = \epsilon_n \frac{i^n}{2\pi} \int_C F(iu) J_n(uR) du. \quad (4.3-10)$$

Так как R есть сравнительно медленно изменяющаяся функция времени, то можно ожидать, что то же справедливо и по отношению к $A_n(R)$, по крайней мере для небольших значений n . Поэтому из (4.3—9) видно, что энергетический спектр I будет состоять из последовательности полос; n -я полоса сосредоточена вокруг частоты nf_m . Если при помощи фильтра устранить все полосы за исключением n -й, то выходной ток будет иметь огибающую $A_n(R)$, когда $n \geq 1$. Положив $n=0$, видим, что выходной ток низкой частоты равен просто

$$A_0(R) = \frac{1}{2\pi} \int_C F(iu) J_0(uR) du. \quad (4.3-11)$$

Положив $n=1$, для полосы частот, сосредоточенных вокруг f_m , найдем

$$\frac{A_1(R)}{R} V. \tag{4.3-12}$$

Статистические свойства выходного тока низкой частоты и огибающих выходных полос могут быть получены из соответствующих статистических свойств R . Например, плотность вероятностей для $A_n(R)$ имеет вид

$$\rho(R) \left| \frac{dA_n(R)}{dR} \right. \tag{4.3-13}$$

где $\rho(R)$ есть плотность вероятностей для R . В этом выражении R рассматривается как функция A_n .

Предполагалось, что учтены все полосы, окружающие гармонические частоты nf_m . Если учтем только часть их, то, повидимому, статистические свойства будут стремиться приблизиться к статистическим свойствам шумового тока в соответствие с первым утверждением, сделанным в начале этого раздела.

Если применить (4.3-11) к квадратичному детектору, то получим

$$F(iu) = \frac{2\alpha}{(iu)^3}$$

$$A_0(R) = -\frac{2\alpha}{2\pi i} \int_{(0+)}^{(0+)} \frac{J_0(uR)}{u^3} du = \frac{\alpha}{2} R^2.$$

Применяя (4.3-11) к линейному детектору, получим

$$F(iu) = -\frac{\alpha}{u^2},$$

$$A_0(R) = -\frac{\alpha}{2\pi} \int_{-\infty}^{+\infty} \frac{J_0(uR)}{u^2} du = \frac{\alpha R}{\pi}$$

где путь интегрирования у начала координат отступает книзу. Эти результаты согласуются с результатами, полученными в разделах 4.1 и 4.2.

В качестве последнего примера найдем выходной ток низкой частоты линейного детектора со смещением, выразив его через огибающую R приложенного напряжения. Из таблицы для $F(iu)$, приведенной в Приложении 4А, видим, что $F(iu)$, соответствующее

$$I=0, \quad V < B,$$

$$I=V-B, \quad V > B,$$

равно

$$F(iu) = -\frac{e^{-iuB}}{u^2}.$$

Соответственно выходной ток низкой частоты равен

$$A_0(R) = -\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-i u B} J_0(u R) u^{-2} du,$$

где путь интегрирования отстает книзу у начала координат. Когда $B > R$, интеграл равен нулю, так как путь интегрирования может быть замкнут на нижней полуплоскости бесконечным полукругом. Это значение также следует из физической сущности вопроса. Если $-R < B < R$, то, интегрируя по частям, получим

$$\begin{aligned} A_0(R) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-i u B} [i B J_0(u R) + R J_1(u R)] u^{-1} du = \\ &= -\frac{B}{2} + \frac{1}{\pi} \int_0^{\infty} [B \sin u B J_0(u R) + R \cos u B J_1(u R)] u^{-1} du = \\ &= -\frac{B}{2} + \frac{B}{\pi} \arcsin \frac{B}{R} + \frac{1}{\pi} \sqrt{R^2 - B^2} = \\ &= -\frac{B}{2} + \frac{R}{\pi} F\left(-\frac{1}{2}, -\frac{1}{2}; \frac{1}{2}; \frac{B^2}{R^2}\right), \quad -R < B < R. \end{aligned} \quad (4.3-14)$$

Эта гипергеометрическая функция встречается опять в уравнении (4.7-6). В пределах $-R < B < R$

$$\frac{dA_0}{dR} = \frac{1}{\pi} \sqrt{1 - \frac{B^2}{R^2}}.$$

Если B отрицательно, а $R < -B$, то путь интегрирования может быть замкнут бесконечным полукругом на верхней полуплоскости и значение интеграла пропорционально вычету полюса в начале координат

$$A_0(R) = 2\pi i \left(-\frac{1}{2\pi}\right) (-iB) = -B.$$

Поэтому выходной ток низкой частоты для линейного детектора равен при $B > 0$ (R всегда положительно)

$$A_0(R) = 0, \quad R < B, \quad (4.3-15)$$

$$A_0(R) = -\frac{B}{2} + \frac{B}{\pi} \arcsin \frac{B}{R} + \frac{1}{\pi} \sqrt{R^2 - B^2}, \quad B < R,$$

а при $B < 0$ равен

$$A_0(R) = |B|, \quad R < |B|, \quad (4.3-16)$$

$$A_0(R) = \frac{|B|}{2} + \frac{|B|}{\pi} \arcsin \frac{|B|}{R} + \frac{1}{\pi} \sqrt{R^2 - B^2}, \quad |B| < R,$$

где значения арксинусов лежат между 0 и $\frac{\pi}{2} \cdot A_0(R)$ и ее первая производная по R непрерывны.

Из (4.3—15) постоянная составляющая выходного тока при $B > 0$ равна

$$I_{dc} = \int_B^{\infty} \left[-\frac{B}{2} + \frac{B}{\pi} \arcsin \frac{B}{R} + \frac{1}{\pi} \sqrt{R^2 - B^2} \right] p(R) dR, \quad (4.3-17)$$

где $p(R)$ есть плотность вероятностей огибающей входного напряжения V , т. е. $p(R)$ имеет вид (3.7—10) для случая одних шумов и вид (3.10—11) для шумов плюс синусоидальное напряжение. Подобным же образом эффективное значение тока низкой частоты (без постоянной составляющей) I_{lf} может быть найдено из

$$\overline{I_{lf}^2} = \overline{I_{fl}^2} - I_{dc}^2,$$

где, если $B > 0$,

$$\overline{I_{fl}^2} = \int_B^{\infty} \left[-\frac{B}{2} + \frac{B}{\pi} \arcsin \frac{B}{R} + \frac{1}{\pi} \sqrt{R^2 - B^2} \right]^2 p(R) dR. \quad (4.3-18)$$

Если V состоит из синусоидального напряжения с амплитудой P плюс шумы V_N , то оно может быть представлено как (4.3—13), а если P значительно больше эффективного значения V_N , то распределение R приблизительно подчиняется нормальному закону. Если, вдобавок, $(P-B) \gg$ (эфф. знач. V_N) > 0 , то (4.3—17), (4.3—18) и (3.10—19) приводят к приближениям

$$\begin{aligned} I_{dc} &\approx -\frac{B}{2} + \frac{B}{\pi} \arcsin \frac{B}{P} + \frac{1}{\pi} \sqrt{P^2 - B^2} + \frac{\psi_0}{2\pi\sqrt{P^2 - B^2}} \approx \\ &\approx -\frac{B}{2} + \frac{P}{\pi} + \frac{B^2 + \psi_0}{2\pi P}, \\ \overline{I_{lf}^2} &\approx \frac{P^2 - B^2}{\pi^2 P^2} \psi_0. \end{aligned} \quad (4.3-19)$$

Второе выражение для I_{dc} предполагает, что $P \gg B$. Если $B=0$, то эти выражения сводятся к первым членам (4.2—5) и (4.2—7). Применяя другой метод, Миддлтон получил более точную формулу для этого случая.

Для данного приложенного напряжения ток $I_{dc}(+)$, соответствующий положительному смещению $|B|$, связан с $I_{dc}(-)$, соответствующему отрицательному смещению $-|B|$, следующим образом:

$$I_{dc}(-) = |B| + I_{dc}(+). \quad (4.3-20)$$

Точно так же эффективное значение $I_{lf}(+)$ равно эффективному значению $I_{lf}(-)$. Уравнение (4.3—20) следует из физических

соображений, основанных на рассмотрении площадей, ограниченных кривой I . Оба приведенных соотношения вытекают из формул, данных Миддльтоном для случая, когда V состоит из синусоидального напряжения и шумов. Они могут быть также выведены из (4.3—17) и (4.3—18).

4.4. ВЫХОДНОЙ ЭНЕРГЕТИЧЕСКИЙ СПЕКТР

Рассмотрим теперь методы решения следующей задачи: дано нелинейное устройство и входное напряжение, состоящее либо из одних шумов, либо из шумов и сигнала. Каков будет выходной энергетический спектр?

В некоторых отношениях ответ на этот вопрос дает меньше полезных сведений, нежели методы исследования, рассмотренные в первых трех разделах. Например, помимо определения эффективного значения, он говорит очень мало о плотности вероятностей тока, соответствующего данной полосе частот на выходе. С другой стороны это эффективное значение может быть найдено (путем интегрирования энергетического спектра) для любой исследуемой полосы частот. Описанные ранее методы ограничены случаем, когда входное напряжение занимает сравнительно узкую полосу частот. Даваемые ими сведения относятся ко всей полосе, соответствующей данной гармонике (0-й, 1-й, 2-й и т. д.) входного напряжения. Не существует другого пути для изучения выходного эффекта, когда часть полосы частот задержана фильтрами, за исключением нахождения энергетического спектра некоторой функции огибающей.

В настоящее время имеются два общих метода, пригодных для определения выходного энергетического спектра, причем каждый из них обладает своими преимуществами и недостатками. При прямом методе шумы представляются в виде суммы конечного числа синусоидальных составляющих и вычисляются типичные комбинационные тона. Выходной энергетический спектр получается из рассмотрения плотностей и амплитуд этих комбинационных тонов. Главное преимущество этого метода заключается в его тесной связи с известной теорией искажений в нелинейных системах. Вообще комбинационные тона низшего порядка являются единственными, имеющими существенное значение для мощности на выходе, и если они известны, то вопрос близок к решению. Основной недостаток метода — трудность подсчета комбинационных тонов, приходящихся на данный интервал. Однако Беннет разработал способ решения этого вопроса ¹⁾.

Основная идея второго метода заключается в нахождении функции корреляции для выходного тока. Отсюда выходной энергетический спектр может быть получен путем преобразования Фурье.

¹⁾ *BSTJ*, 19, 587—610, Приложение В, 1940.

Метод функции корреляции и его варианты имеют более недавнее происхождение, нежели прямой метод.

В методе функции корреляции обойден вопрос о подсчете комбинационных тонов. Однако в некоторых случаях он становится несколько ограниченным. Вероятно, лучше всего при исследовании какого-либо вопроса иметь в виду оба метода. Прямой метод будет иллюстрирован применением его к квадратичному детектору. В дальнейшем будут приведены примеры двух решений методом функций корреляции.

4.5. ЭНЕРГЕТИЧЕСКИЙ СПЕКТР НА ВЫХОДЕ УСТРОЙСТВА С КВАДРАТИЧНОЙ ХАРАКТЕРИСТИКОЙ

Вероятно, наиболее прямой путь получения $W(f)$ — энергетического спектра тока I , где

$$I = \alpha V^2, \quad (4.1-1)$$

а V — напряжение шумов, заключается в возведении в квадрат выражения

$$V = V_N = \sum_1^M c_m \cos(\omega_m t - \varphi_m), \quad (2.8-6)$$

в котором $c_m^2 = 2\omega(f_m)\Delta f$, $\omega_m = 2\pi f_m$, $f_m = m\Delta f$, а $\varphi_1, \varphi_2, \dots, \varphi_M$ — случайные фазовые углы.

Значительное упрощение алгебраических выкладок достигается при замене (2.8-6) на

$$V_N = \frac{1}{2} \sum_{-\infty}^{+\infty} c_m e^{imat - i\varphi_m}. \quad (4.5-1)$$

Здесь добавлен член $c_0/2$, чтобы не было пробелов в суммировании, и введены обозначения

$$\begin{aligned} c_{-m} &= c_m, \\ \varphi_{-m} &= -\varphi_m, \\ a &= 2\pi\Delta f. \end{aligned} \quad (4.5-2)$$

Возведение в квадрат (4.5-1) дает двойные ряды

$$\begin{aligned} V_N^2 &= \frac{1}{4} \sum_{-\infty}^{+\infty} \sum_{-\infty}^{+\infty} c_m c_n e^{i(m+n)at - i\varphi_m - i\varphi_n} = \\ &= \frac{1}{4} \sum_{k=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} c_{k-n} c_n e^{ikat - i\varphi_{k-n} - i\varphi_n}. \end{aligned}$$

Допустим, что надо рассмотреть составляющую V_N^2 частоты $f_k = k\Delta f$. Очевидно, она равна

$$A_k \cos(\omega_k t - \psi_k) = \frac{1}{2} \sum_{n=-\infty}^{+\infty} c_{k-n} c_n \cos(kat - \varphi_{k-n} - \varphi_n). \quad (4.5-3)$$

$W(f_k)$ — энергетический спектр тока I на частоте f_k равен α^2 , умноженному на коэффициент при Δf в среднем квадрате (4.5-3), где усреднение производится по φ .

Поэтому

$$W(f_k) \Delta f = \frac{\alpha^2}{4} \sum_{-\infty}^{+\infty} \sum_{-\infty}^{+\infty} c_{k-n} c_n c_{k-m} c_m \times \\ \times \left[\overline{\cos(kat - \varphi_{k-n} - \varphi_n) \cos(kat - \varphi_{k-m} - \varphi_m)} \right],$$

где суммирование распространяется на m и n . Пусть n — фиксировано; рассмотрим те значения m , которые дают среднее, отличное от нуля. Видно, что $m=n$ и $m=k-n$ суть два таких значения. Единственные другие возможности суть $m=-n$ и $m=-k+n$, но они приводят к членам, содержащим (за исключением случаев, когда n или k равно нулю) три различных угла φ_n , φ_{k-n} и φ_{k+n} , усреднение по которым дает нуль. Используя то обстоятельство, что среднее косинусов, возводимых в квадрат, есть $\frac{1}{2}$ и что для данного n имеются два таких члена, получим

$$W(f_k) \Delta f = \frac{\alpha^2}{4} \sum_{n=-\infty}^{+\infty} c_{k-n}^2 c_n^2 = \alpha^2 \Delta f \sum_{n=-\infty}^{+\infty} \omega(f_k - f_n) \omega(f_n) \Delta f, \quad (4.5-4)$$

где в последнем выражении было использовано

$$f_{k-n} = (k-n)\Delta f = f_k - f_n$$

и учтено, что из $c_{-n} = c_n$ следует

$$\omega(f_{-n}) = \omega(-n \Delta f) = \omega(-f_n) = \omega(f_n).$$

Поэтому из (4.5-4) найдем для энергетического спектра тока I

$$W(f) = \alpha^2 \int_{-\infty}^{+\infty} \omega(x) \omega(f-x) dx, \quad (4.5-5)$$

подразумевая, что $f \neq 0$ и что

$$\omega(-x) = \omega(x). \quad (4.5-6)$$

Результат, который получается при использовании (2.8—6) с введением косинусов и только при положительных значениях m , равен

$$W(f) = \alpha^2 \int_0^f \omega(x) \omega(f-x) dx + 2\alpha^2 \int_0^\infty \omega(x) \omega(f+x) dx. \quad (4.5-7)$$

Он содержит только положительные значения частоты. Уравнения (4.5—5) и (4.5—7) эквивалентны и легко могут быть преобразованы одно в другое.

Первый интеграл в (4.5—7) учитывает суммарные комбинационные тона второго порядка, а второй интеграл — разностные комбинационные тона. Это можно показать путем написания тока в виде

$$\begin{aligned} I &= \alpha V^2 = \alpha \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} c_m c_n \cos(\omega_m t - \varphi_m) \cos(\omega_n t - \varphi_n) = \\ &= \frac{\alpha}{2} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} c_m c_n \{ \cos[(\omega_m - \omega_n)t - \varphi_m + \varphi_n] + \\ &\quad + \cos[(\omega_m + \omega_n)t + \varphi_m + \varphi_n] \}. \end{aligned} \quad (4.5-8)$$

Мощность в полосе частот ($f_k, f_k + \Delta f$) состоит из мощности, создаваемой разностными комбинационными тонами ($\omega_{k+l} - \omega_l$), плюс мощность, создаваемая суммарными комбинационными тонами ($\omega_{k-l} + \omega_l$). В первом случае l пробегает от 1 до ∞ , а во втором случае l пробегает от 1 до $k-1$.

Рассмотрим сначала разностные комбинационные тона и на момент предположим, что как k , так и l фиксированы. Два ряда значений $m=k+l, n=l$ и $m=l, n=k+l$ суть единственные значения m и n в (4.5—8), дающие частоты ($\omega_{k+l} - \omega_l$). Два соответствующих члена в (4.5—8) равны, так как $\cos(-x)$ равен $\cos x$. Средняя мощность, связанная с этими двумя членами,

$$\begin{aligned} \left(\frac{\alpha}{2} c_{k+l} c_l \right)^2 \overline{4 \cos^2[(\omega_{k+l} - \omega_l)t - \varphi_{k+l} + \varphi_l]} = \\ = \frac{1}{2} (\alpha c_{k+l} c_l)^2. \end{aligned} \quad (4.5-9)$$

Мощность в полосе частот ($f_k, f_k + \Delta f$), связанная с разностными тонами, получается суммированием по l от 1 до ∞ :

$$\frac{\alpha^2}{2} \sum_{l=1}^{\infty} c_{k+l}^2 c_l^2 = 2\alpha^2 \sum_{l=1}^{\infty} \omega(f_{k+l}) \omega(f_l) (\Delta f)^2 \rightarrow 2\alpha^2 \Delta f \int_0^\infty \omega(f_k + f) \omega(f) df.$$

Это приводит ко второму члену в (4.5—7).

Рассмотрим теперь суммарные комбинационные тона. Для членов такого типа в (4.5—8), приводящих к частоте ω_k , $m+n=k$. Пусть $n=1$, тогда $m=k-1$. Фаза этого члена по отношению ко всем другим членам случайна, за исключением одного члена, соответствующего $n=k-1$, $m=1$, который имеет такую же фазу. Средняя мощность, вносимая этими двумя членами в (4.5—8), равна, как и в (4.5—9),

$$\frac{1}{2}(\alpha c_1 c_{k-1})^2.$$

Она связана с двумя членами, для которых $m+n=k$. Положив $n=2$ и проделав ту же процедуру, получим еще два члена. Поэтому, предположив, что k — нечетное число, найдем, что мощность, создаваемая в интервале $(f_k, f_k + \Delta f)$ суммарными комбинационными тонами, равна

$$\frac{1}{2} \sum_{n=1}^{(k-1)/2} (\alpha c_n c_{k-n})^2 = \frac{1}{4} \sum_{n=1}^{k-1} (\alpha c_n c_{k-n})^2 \rightarrow \alpha^2 \Delta f \int_0^{f_k} \omega(f) \omega(f_k - f) df,$$

что приводит к первому члену в (4.5—7).

Если напряжение V , приложенное к устройству с квадратичной характеристикой, является суммой напряжения шумов V_N и синусоидального напряжения

$$V = P \cos pt + V_N, \quad (4.1-13)$$

то получим

$$V^2 = P^2 \cos^2 pt + 2PV_N \cos pt + V_N^2. \quad (4.5-10)$$

Из двух уравнений

$$\cos^2 pt = \frac{1}{2} + \frac{1}{2} \cos 2pt,$$

$$\overline{V_N^2} = \sum_1^M c_m^2 \frac{1}{2} \rightarrow \int_0^\infty \omega(f) df$$

следует, что $I = \alpha V^2$ имеет постоянную составляющую

$$\frac{\alpha P^2}{2} + \alpha \int_0^\infty \omega(f) df, \quad (4.5-11)$$

что согласуется с (4.1—14), и гармоническую составляющую

$$\frac{\alpha P^2}{2} \cos 2pt. \quad (4.5-12)$$

Сплошной участок энергетического спектра $W_c(f)$ определяется остальными членами (4.5—10)

$$2PV_N \cos pt + V_N^2.$$

Используя (2.8—6), видим, что

$$2PV_N \cos pt = P \sum_1^M c_m [\cos(\omega_m t + pt - \varphi_m) + \cos(\omega_m t - pt - \varphi_m)].$$

Предположим, что $p = 2\pi r \Delta f$. Члены, имеющие отношение к частоте $f_n = n \Delta f$, это те, для которых

$$\begin{aligned} \omega_m + p &= 2\pi f_n, & |\omega_m - p| &= 2\pi f_n, \\ m + r &= n, & |m - r| &= n, \\ m &= n - r, & m &= r \pm n, \end{aligned}$$

где должны быть приняты во внимание только положительные значения m . Если $n > r$, то $m = n - r$ или $m = r + n$. Если $n < r$, то $m = r - n$ или $m = r + n$. В обоих случаях значениями m являются $|n - r|$ или $n + r$. Члены частоты f_n в выражении для $2PV_n \cos pt$ поэтому равны

$$Pc_{|n-r|} \cos(2\pi f_n t - \varphi_{|n-r|}) + Pc_{n+r} \cos(2\pi f_n t - \varphi_{n+r}),$$

а средний квадрат этого выражения, если усреднение взять по φ , равен

$$\begin{aligned} \frac{P^2}{2} (c_{|n-r|}^2 + c_{n+r}^2) &= P^2 \Delta f [\omega(f_{|n-r|}) + \omega(f_{n+r})] = \\ &= P^2 \Delta f [\omega(|f_n - f_p|) + \omega(f_n + f_p)], \end{aligned}$$

где $f_p = p/2\pi$.

Добавляя еще выражение (4.5—5), полученное из V_N^2 , видим, что сплошной участок $W_c(f)$ энергетического спектра тока I равен

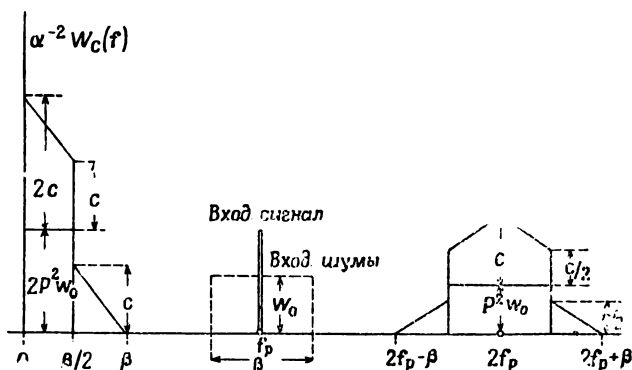
$$\begin{aligned} W_c(f) &= \alpha^2 P^2 [\omega(f - f_p) + \omega(f + f_p)] + \\ &+ \alpha^2 \int_{-\infty}^{+\infty} \omega(x) \omega(f - x) dx, \end{aligned} \quad (4.5-13)$$

где $\omega(-f)$ имеет такое же значение, что и $\omega(f)$.

Уравнение (4.5—13) было использовано для вычисления $W_c(f)$ (фиг. 8). Предполагается, что составляющие входных шумов имеют одинаковые амплитуды в полосе частот шириной β с центральной частотой f_p (ср. фильтр c , Приложение 4С). Рассматривая площадь, ограниченную кривой низкочастотной части спектра, найдем

$$\int_0^{\beta} W_c(f) df = \alpha^2 \beta \omega_0 (P^2 + \beta \omega_0).$$

Так как средний квадрат входного напряжения V_N есть $\psi_0 = \beta \omega_0$, то это уравнение совпадает с выражением (4.1—15) для среднего квадрата тока I_{lf} (ток низкой частоты без постоянной составляющей). Если фильтры звуковых частот срезают часть спектра, то, чтобы получить средний квадрат соответствующего выходного тока, $W_c(f)$ следует интегрировать по сохранившейся части спектра, о чем упоминается в примечании к уравнению (4.1—6).



Фиг. 8. Сплошной участок спектра на выходе устройства с квадратичной характеристикой.

Напряжение на входе $= P \cos 2\pi f_p t + V_N$. Постоянная составляющая на выходе $= \alpha \left(\frac{P^2}{2} + \beta \omega_0 \right) \beta \omega_0^2 = C$.

Когда напряжение V состоит из напряжения шумов V_N и двух синусоидальных напряжений, частоты которых не являются кратными:

$$V = P \cos pt + Q \cos qt + V_N,$$

то можно показать, что сплошной участок $W_c(f)$ энергетического спектра тока I равен (4.5—13) плюс дополнительные члены

$$\alpha^2 Q^2 [\omega(f - f_q) + \omega(f + f_q)], \quad (4.5-14)$$

где $f_q = q/2\pi$.

Если напряжение, приложенное к устройству с квадратичной характеристикой (4.1—1), равно

$$\begin{aligned} V(t) &= Q(1 + k \cos pt) \cos qt + V_N = \\ &= Q \cos qt + \frac{Qk}{2} \cos(p+q)t + \frac{Qk}{2} \cos(p-q)t + V_N, \end{aligned}$$

то результирующий ток содержит постоянную составляющую

$$\frac{\alpha}{2} Q^2 \left(1 + \frac{k^2}{2} \right) + \alpha \int_0^{\infty} \omega(f) df. \quad (4.5-15)$$

Синусоидальные члены тока I можно получить возведением в квадрат

$$Q(1 + k \cos pt) \cos qt$$

и умножением на α . Оставшаяся часть тока I имеет сплошной энергетический спектр

$$\begin{aligned} W_c(f) = \alpha^2 Q^2 & \left[\omega(f - f_q) + \omega(f + f_q) + \right. \\ & + \frac{k^2}{4} \omega(f - f_p - f_q) + \frac{k^2}{4} \omega(f + f_p + f_q) + \\ & \left. + \frac{k^2}{4} \omega(f - f_p + f_q) + \frac{k^2}{4} \omega(f + f_p - f_q) \right] + \\ & + \alpha^2 \int_{-\infty}^{+\infty} \omega(x) \omega(f - x) dx, \end{aligned} \quad (4.5-16)$$

где $f_p = p/2\pi$ и $f_q = q/2\pi$.

4.6. ДВА МЕТОДА ФУНКЦИИ КОРРЕЛЯЦИИ

Как упоминалось в разделе 4.4, эти методы определения выходного энергетического спектра основаны на нахождении функции корреляции $\Psi(\tau)$ для выходного тока. Отсюда энергетический спектр $W(f)$ выходного тока можно определить из (2.1-5), переписанного в виде

$$W(f) = 4 \int_0^{\infty} \Psi(\tau) \cos 2\pi f\tau d\tau. \quad (4.6-1)$$

Напомним, что $W(f)\Delta f$ может рассматриваться как средняя мощность, которая рассеивалась бы составляющими тока I , лежащими в полосе частот $(f, f + \Delta f)$, если I протекает через сопротивление 1 ом.

Допустим, что на вход нелинейного устройства подано напряжение $V(t)$. Оно может состоять из напряжения шумов $V_N(t)$ плюс синусоидальные составляющие. Предположим, что ток на выходе есть $I(t)$. Нелинейное устройство характеризуется связью между $V(t)$ и $I(t)$. В данной работе предполагается, что $I(t)$ в момент t полностью определяется значением $V(t)$ в момент t .

Будут описаны два метода нахождения $\Psi(\tau)$:

а) интегрированием двумерной плотности вероятностей для $V(t)$ и $V(t + \tau)$ в пределах, заданных свойствами нелинейного устройства; этот метод является наиболее прямым в случае, когда на детектор действуют только одни шумы;

б) введением и использованием характеристической функции двумерной плотности вероятностей для $V(t)$ и $V(t+\tau)$; ради краткости характеристическая функция будет обозначаться x . ф.

4.7. ЛИНЕЙНОЕ ДЕТЕКТИРОВАНИЕ ШУМОВ — ПЕРВЫЙ МЕТОД

Первый метод исследования будет иллюстрирован применением его к определению энергетического спектра на выходе линейного детектора, когда входное напряжение состоит только из одних шумов.

Линейный детектор характеризуется следующими соотношениями:

$$I(t) = \begin{cases} 0, & V(t) < 0, \\ V(t), & V(t) > 0, \end{cases} \quad (4.7-1)$$

которые могут быть получены из (4.2—1), если положить $\alpha=1$, а входное напряжение равным

$$V(t) = V_N(t), \quad (4.7-2)$$

где $V_N(t)$ — шумовое напряжение, функция корреляции которого есть $\psi(\tau)$, а энергетический спектр $\omega(f)$.

Функция корреляции $\Psi(\tau)$ есть среднее значение $I(t)I(t+\tau)$, которое равно среднему значению функции

$$F(V_1, V_2) = \begin{cases} V_1 V_2, & \text{когда } V_1, V_2 > 0, \\ 0, & \text{при всех других } V, \end{cases} \quad (4.7-3)$$

где было положено

$$\begin{aligned} V_1 &= V(t), \\ V_2 &= V(t + \tau). \end{aligned}$$

Двумерное распределение для V_1 и V_2 находится из (3.2—4), откуда следует, что среднее значение функции $F(V_1, V_2)$ есть

$$\int_{-\infty}^{+\infty} dV_1 \int_{-\infty}^{+\infty} dV_2 \frac{F(V_1, V_2)}{2\pi |M|^{1/2}} \exp \left[-\frac{1}{2|M|} (\psi_0 V_1^2 + \psi_0 V_2^2 - 2\psi_\tau V_1 V_2) \right], \quad (4.7-4)$$

где

$$|M| = \psi_0^2 - \psi_\tau^2.$$

Для случая линейного детектора, когда $F(V_1, V_2)$ находится из (4.7—3), интеграл равен

$$|M|^{-1/2} \frac{1}{2\pi} \int_0^{\infty} dV_1 \int_0^{\infty} dV_2 V_1 V_2 \exp \left[-\frac{1}{2|M|} (\psi_0 V_1^2 + \psi_0 V_2^2 - 2\psi_\tau V_1 V_2) \right] =$$

$$= \frac{1}{2\pi} \left[(\psi_0^2 - \psi_\tau^2)^{1/2} + \psi_\tau \arccos \left(\frac{-\psi_\tau}{\psi_0} \right) \right],$$

где для вычисления интеграла было использовано (3.5—4). Значение арккосинуса берется в пределах от 0 до π .

Итак, функция корреляции для $I(t)$ равна

$$\Psi(\tau) = \frac{1}{2\pi} \left[(\psi_0^2 - \psi_\tau^2)^{1/2} + \psi_\tau \arccos \left(\frac{-\psi_\tau}{\psi_0} \right) \right]. \quad (4.7-5)$$

Отсюда при помощи (4.6—1) может быть получен энергетический спектр $W(f)$. С этой целью удобно написать (4.7—5) через гипергеометрическую функцию. Путем разложения и сравнения членов найдем

$$\begin{aligned} \Psi(\tau) &= \frac{\psi_\tau}{4} + \frac{\psi_0}{2\pi} F \left(-\frac{1}{2}, -\frac{1}{2}; \frac{1}{2}; \frac{\psi_\tau^2}{\psi_0^2} \right) = \\ &= \frac{\psi_\tau}{4} + \frac{\psi_0}{2\pi} + \frac{\psi_\tau^2}{4\pi\psi_0} + \text{члены с } \psi_\tau^4, \psi_\tau^6 \text{ и т. д.} \end{aligned} \quad (4.7-6)$$

Как будет более подробно изложено в разделе 4.8, постоянный член A^2 в $\Psi(\tau)$ относится к постоянной составляющей тока $I(t)$ силой в A ампер. Поэтому $I(t)$ имеет постоянную составляющую, равную

$$\left[\frac{\psi_0}{2\pi} \right]^{1/2} = \frac{1}{\sqrt{2\pi}} \times [\text{эфф. знач. } V(t)]. \quad (4.7-7)$$

Это совпадает с (4.2—3), если положить P равным нулю.

Интегралы вида

$$G_n(f) = \int_0^\infty \psi_\tau^n \cos 2\pi f\tau \, d\tau,$$

которые получаются, если (4.7—6) подставить в (4.6—1) и почленно интегрировать, рассматриваются в Приложении 4С. Из приведенных там результатов видно, что если пренебречь ψ_τ^4 и высшими степенями, то получим следующее приближение для сплошного участка $W_c(f)$ спектра $W(f)$:

$$\begin{aligned} W_c(f) &\approx G_1(f) + \frac{G_2(f)}{\pi\psi_0} = \\ &= \frac{\omega(f)}{4} + \frac{1}{4\pi\psi_0} \frac{1}{2} \int_{-\infty}^{+\infty} \omega(x)\omega(f-x) \, dx, \end{aligned} \quad (4.7-8)$$

где $\omega(-f)$ определено как $\omega(f)$.

Если спектр $V_N(f)$ равномерен в сравнительно узком диапазоне, простирающемся от f_a до f_b , так что $\omega(f)$ равен ω_0 в этом диапазоне и нулю вне его, то можно воспользоваться полученными в Приложении 4С результатами для фильтра с. Приведенные там величины f_0 и β связаны с f_a и f_b следующим образом:

$$f_a = f_0 - \frac{\beta}{2}, \quad f_b = f_0 + \frac{\beta}{2},$$

а значение ω_0 такое же, что и в данном случае, и равно $\frac{\psi_0}{\beta}$. Приведенное там значение $G_2(f)$ ведет к следующему приближению для низких частот:

$$W_c(f) \approx \frac{1}{\pi\psi_0} \frac{\psi_0^2}{4\beta} \left(1 - \frac{f}{\beta}\right) = \frac{\omega_0}{4\pi} \left(1 - \frac{f}{f_b - f_a}\right), \quad (4.7-9)$$

когда $0 < f < (f_b - f_a)$, и к $W_c(f) \approx 0$ для $(f_b - f_a) < f < f_a$.

Положив для кривой $W_c(f)$ (фиг. 8), соответствующей квадратичному детектору, $P=0$, видим, что форма низкочастотного участка спектра представляет собой треугольник, а при $f=\beta$ спектр равен нулю. Поэтому из (4.7-9) можно прийти к заключению, что в первом приближении форма выходного энергетического спектра в случае линейного детектора такова же, что и для квадратичного детектора, если входное напряжение представляет собой шумы в сравнительно узкой полосе частот.

Приближенное эффективное значение низкочастотного выходного тока может быть получено интегрированием (4.7-9)

$$\overline{I_{if}^2} = \int_0^{f_b - f_a} W_c(f) df \approx \frac{\omega_0(f_b - f_a)}{8\pi} = \frac{\psi_0}{8\pi},$$

откуда

$$\text{эфф. знач. тока н.ч.} \approx \frac{1}{\sqrt{8\pi}} \times [\text{эфф. знач. прилож. напряжения}]. \quad (4.7-10)$$

Как видно, этот ток равен половине постоянной составляющей. Следует помнить, что (4.7-10) является приближением, так как мы пренебрегли ψ_c^4 и высшими степенями. Точное значение может быть получено из (4.2-8). Тогда коэффициент $(8\pi)^{-1/2} = 0,200$ следует заменить на

$$\frac{1}{\pi} \left(2 - \frac{\pi}{2}\right)^{1/2} = 0,209.$$

Для других типов полосовых фильтров можно найти $W_c(f)$, если из Приложения 4С взять соответствующие значения G . Ока-

зывается, что (4.7—10) справедливо для всех трех типов фильтров. Это частный случай теоремы, упоминаемой несколько раз раньше, которая утверждает, что полная мощность любых комбинационных частот зависит только от полной мощности действующих на входе шумов, но не от их спектрального распределения. Позже в разделе 4.9 будет показано, что член ψ_τ^n в (4.7—6) соответствует комбинационным частотам n -го порядка.

4.8. МЕТОД ХАРАКТЕРИСТИЧЕСКОЙ ФУНКЦИИ

Как упоминалось раньше, особенно в связи с уравнением (1.4—3), х. ф. случайной переменной x есть среднее значение $\exp(iux)$. Она является функцией u . Х. ф. двух случайных переменных x и y есть среднее значение $\exp(iux + ivy)$, т. е. функция u и v .

Х. ф., которой мы будем здесь пользоваться, есть х. ф. двух случайных переменных $V(t)$ и $V(t+\tau)$, где $V(t)$ — напряжение, приложенное к нелинейному устройству. Его случайный характер определяется тем, что t выбирается случайным, а τ остается фиксированным. Эту характеристическую функцию можно написать в виде

$$[g(u, v, \tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \exp[iuV(t) + ivV(t+\tau)] dt. \quad (4.8-1)$$

Если $V(t)$ содержит напряжение шумов $V_N(t)$, как всегда будет предполагаться в этом разделе, и если используются уравнения (2.8—1) или (2.8—6), то в уравнении (4.8—1) появляется большое число случайных параметров (a_n и b_n или φ_n). В соответствии с применением упомянутых уравнений можно произвести усреднение по этим параметрам без изменения значения (4.8—1) и тем самым упростить интегрирование.

Например, положим

$$V(t) = V_s(t) + V_N(t), \quad (4.8-2)$$

где $V_s(t)$ есть некоторое регулярное напряжение, которое может, например, состоять из одной или более синусоидальных составляющих. Подставляя в (4.8—1) и используя соотношение (3.2—7), для характеристической функции $V_N(t)$ и $V_N(t+\tau)$ найдем

$$\begin{aligned} g_N(u, v, \tau) &= \text{сред.} \left\{ \exp \left[iuV_N(t) + ivV_N(t + \tau) \right] \right\} = \\ &= \exp \left[-\frac{\psi_0}{2} (u^2 + v^2) - \psi_\tau uv \right], \end{aligned} \quad (4.8-3)$$

где $\psi_\tau \equiv \psi(\tau)$ — функция корреляции напряжения шумов $V_N(t)$. Для х. ф. $V(t)$ и $V(t+\tau)$ получим

$$g(u, v, \tau) = \exp \left[-\frac{\psi_0}{2} (u^2 + v^2) - \psi_\tau uv \right] \times$$

$$\times \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \exp \left[iuV_s(t) + ivV_s(t + \tau) \right] dt = g_N(u, v, \tau) g_s(u, v, \tau). \quad (4.8-4)$$

В последней строчке $g_s(u, v, \tau)$ обозначает предел выражения в средней строчке

$$g_s(u, v, \tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \exp [iuV_s(t) + ivV_s(t + \tau)] dt. \quad (4.8-5)$$

Основной причиной, по которой удобно пользоваться х. ф., является то, что совсем немногие нелинейные устройства могут быть описаны интегралом

$$I = \frac{1}{2\pi} \int_C F(iu) e^{iVu} du, \quad (4A-1)$$

где функция $F(iu)$ и путь интегрирования C выбираются в соответствии с типом нелинейного устройства. Примеры таких устройств даны в Приложении 4А.

Функция корреляции $\Psi(\tau)$ для $I(t)$ равна

$$\begin{aligned} \Psi(\tau) &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T I(t) I(t + \tau) dt = \\ &= \lim_{T \rightarrow \infty} \frac{1}{4\pi^2 T} \int_0^T dt \int_C F(iu) e^{iuV(t)} du \int_C F(iv) e^{ivV(t+\tau)} dv = \\ &= \frac{1}{4\pi^2} \int_C F(iu) du \int_C F(iv) dv \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \exp [iuV(t) + ivV(t + \tau)] dt = \\ &= \frac{1}{4\pi^2} \int_C F(iu) du \int_C F(iv) g(u, v, \tau) dv. \end{aligned} \quad (4.8-6)$$

Это основная формула метода характеристической функции.

Если $V(t)$ есть сумма напряжения шумов и регулярного напряжения, как (4.8-2), то (4.8-6) превращается в

$$\Psi(\tau) \approx \frac{1}{4\pi^2} \int_C F(iu) e^{-(\psi_0/2)u^2} du \int_C F(iv) e^{-(\psi_0/2)v^2} e^{-\frac{\psi_\tau uv}{\tau}} g_s(u, v, \tau) dv, \quad (4.8-7)$$

где $g_s(u, v, \tau)$ есть х. ф. для $V_s(t)$ и $V_s(t + \tau)$, определяемая уравнением (4.8—5). Это окончательное выражение для $\Psi(\tau)$. Все дальнейшее сводится к вычислению этого интеграла и к нахождению из уравнения

$$W(f) = 4 \int_0^{\infty} \Psi(\tau) \cos 2\pi f \tau d\tau \quad (4.6-1)$$

энергетического спектра для I .

Весьма часто ток $I(t)$ имеет постоянную и периодические составляющие. Представляется удобным рассматривать последние отдельно, так как они соответствуют тем членам в $\Psi(\tau)$, благодаря которым интеграл (4.6—1) для $W(f)$ расходится. Действительно, из раздела 2.2 следует, что функция корреляции вида

$$A^2 + \frac{C^2}{2} \cos 2\pi f_0 \tau \quad (2.2-3)$$

соответствует току

$$A + C \cos(2\pi f_0 t - \varphi), \quad (2.2-2)$$

где фазовый угол φ не может быть определен из (2.2—3), так как он не влияет на среднюю мощность.

Рассмотрим функцию корреляции для $V(t) = V_s(t) + V_N(t)$. Она равна

$$\begin{aligned} \lim_{T \rightarrow \infty} \frac{1}{T} & \left[\int_0^T V_s(t) V_s(t + \tau) dt + \int_0^T V_s(t) V_N(t + \tau) dt + \right. \\ & \left. + \int_0^T V_N(t) V_s(t + \tau) dt + \int_0^T V_N(t) V_N(t + \tau) dt \right]. \quad (4.8-8) \end{aligned}$$

Так как $V_s(t)$ и $V_N(t)$ независимы, то второй и третий интегралы исчезают.

В результате

$$\begin{aligned} & \text{Функция корреляции для } V(t) = \text{функция корреляции для} \\ & V_s(t) + \text{функция корреляции для } V_N(t). \quad (4.3-9) \end{aligned}$$

Теперь, когда $\tau \rightarrow \infty$, функция корреляции для $V_N(t)$ становится равной нулю, а функция корреляции для $V_s(t)$ приобретает приведенную выше форму (2.2—3). Таким образом, функция корреляции регулярного напряжения $V_s(t)$ может быть найдена из $V(t)$, если положить $\tau \rightarrow \infty$ и собрать остающиеся члены.

Результаты, полученные для $V(t)$, справедливы и для $I(t)$, и такая же процедура может быть повторена при собирании тех частей $\Psi(\tau)$, которые соответствуют постоянной и периодическим составляющим $I(t)$.

Обращаясь теперь к (4.8—7), видим, что, когда $\tau \rightarrow \infty$, $\psi_\tau \rightarrow 0$, тогда как $g_s(u, v, \tau)$, определяемая уравнением (4.8—5) и соответствующая напряжению $V_s(t)$, остается, в общем, по величине неизменной. Это последнее утверждение может показаться не совсем очевидным, но исследование ряда случаев позже докажет, что оно справедливо, по крайней мере, для этих случаев. Поэтому, полагая в (4.8—7) $\psi_\tau = 0$, находим ту часть $\Psi(\tau)$, которая соответствует постоянной и периодической составляющим $I(t)$:

$$\Psi_\infty(\tau) = \frac{1}{4\pi^2} \int_C F(iu) e^{-\frac{\psi_0}{2} u^2} du \int_C F(iv) e^{-\frac{\psi_0}{2} v^2} g_s(u, v, \tau) dv, \quad (4.8-10)$$

где индекс ∞ показывает, что $\Psi_\infty(\tau)$ есть та часть $\Psi(\tau)$, которая не исчезает при $\tau \rightarrow \infty$.

Уравнение (4.8—9) в применении к $I(t)$ можно написать в виде

$$\Psi(\tau) = \Psi_\infty(\tau) + \Psi_c(\tau), \quad (4.8-11)$$

где $\Psi_c(\tau)$ есть функция корреляции сплошного участка энергетического спектра тока $I(t)$.

Кстати, если воспользоваться импульсными функциями $\delta(f)$ для интерпретации интеграла в (4.8—6), как это изложено в разделе 2.2, то можно избежать разделения $\Psi(\tau)$ на две части в (4.8—11). Этот метод дает надлежащие значения постоянной и синусоидальной составляющих, если даже (4.6—1) не сходится [вследствие наличия членов, ведущих к $\Psi_\infty(\tau)$].

4.9. ЭНЕРГЕТИЧЕСКИЙ СПЕКТР НА ВЫХОДЕ НЕЛИНЕЙНОГО УСТРОЙСТВА ОБЩЕГО ТИПА

Чтобы иллюстрировать метод характеристической функции, рассмотрим случай нелинейного устройства общего типа, описываемого уравнением

$$I = \frac{1}{2\pi} \int_C F(iu) e^{iVu} du, \quad (4A-1)$$

где V состоит из напряжения шумов плюс синусоидальное напряжение

$$V(t) = P \cos pt + V_N(t). \quad (4.1-13)$$

Как обычно, $V_N(t)$ имеет энергетический спектр $\omega(f)$ и функцию корреляции $\psi(\tau)$. Ради краткости $\psi(\tau)$ часто записывается в виде ψ_τ . Сравнивая (4.1—13) с (4.8—2), найдем

$$V_s(t) = P \cos pt. \quad (4.9-1)$$

Первой задачей является вычисление характеристической функции $g_s(u, v, \tau)$ для двух случайных переменных $V_s(t)$ и $V_s(t+\tau)$. Это можно сделать, пользуясь интегралом (4.8—5):

$$g_s(u, v, \tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \exp [iuP \cos pt + ivP \cos p(t + \tau)] dt = \\ = J_0(P\sqrt{u^2 + v^2 + 2uv \cos p\tau}), \quad (4.9-2)$$

где J_0 — функция Бесселя. Интегрирование выполнено при помощи следующей подстановки:

$$u \cos pt + v \cos p(t + \tau) = (u + v \cos p\tau) \cos pt - v \sin p\tau \sin pt = \\ = \sqrt{u^2 + v^2 + 2uv \cos p\tau} \cos(pt + \text{фазов. угол}),$$

и пользуясь соотношением

$$J_0(z) = \frac{1}{2\pi} \int_0^{2\pi} e^{iz \cos t} dt.$$

Характеристическая функция для (4.1—13) была найдена в разделе 3.10.

Функция корреляции $\Psi(\tau)$ для $I(t)$ теперь может быть получена путем подстановки найденных выше выражений в (4.8—7)

$$\Psi(\tau) = \frac{1}{4\pi^2} \int_C du F(iu) e^{-(\psi_0/2)u^2} \int_C dv F(iv) e^{-(\psi_0/2)v^2} \times \\ \times e^{-\psi_\tau uv} J_0(P\sqrt{u^2 + v^2 + 2uv \cos p\tau}). \quad (4.9-3)$$

$\Psi_\infty(\tau)$ — функция корреляции для постоянной и периодической составляющих I согласно (4.8—10) определяется из этого уравнения, если положить $\psi(\tau) = 0$.

▲ Если надо рассмотреть какой-либо частный случай нелинейного устройства, то соответствующая функция $F(iu)$ может быть найдена в Приложении 4А. Так, например, $F(iu)$ для линейного детектора есть $-u^{-2}$. Подстановка этого значения в (4.9—3) приводит к некоторому двойному интегралу. Если бы существовал какой-либо легкий способ вычисления этого интеграла, то тогда все было бы решено. К сожалению, до сих пор не найдено простого метода вычисления.

Здесь может быть применен один способ, близко связанный с прямым методом. Он основан на разложении

$$g_s(u, v, \tau) = J_0(P\sqrt{u^2 + v^2 + 2uv \cos p\tau}) = \\ = \sum_{n=0}^{\infty} \epsilon_n (-)^n J_n(Pu) J_n(Pv) \cos np\tau; \quad (4.9-4) \\ \epsilon_0 = 1, \quad \epsilon_n = 2 \quad \text{при } n \geq 1$$

Это разложение дает возможность написать те члены в (4.9—3), которые, собственно, создают трудности, в виде

$$e^{-\psi_\tau uv} J_0(P\sqrt{u^2 + v^2 + 2uv \cos p\tau}) = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} (-1)^{n+k} \varepsilon_n \cos n p \tau \frac{(\psi_\tau uv)^k}{k!} J_n(Pu) J_n(Pv). \quad (4.9-5)$$

Преимуществом применения двойной суммы является упрощение интегрирования. Подставляя ее в (4.9—3) и полагая

$$h_{nk} = \frac{i^{n+k}}{2\pi} \int_C F(iu) u^k J_n(Pu) e^{-(\psi_0/2)u^2} du, \quad (4.9-6)$$

получаем

$$\Psi(\tau) = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \frac{1}{k!} \psi_\tau^k h_{nk}^2 \varepsilon_n \cos n p \tau. \quad (4.9-7)$$

Функцию корреляции $\Psi_\infty(\tau)$ для постоянной и периодической составляющих I найдем, полагая $\tau \rightarrow \infty$ и $\psi_\tau \rightarrow 0$. Остаются только те члены, для которых $k=0$:

$$\Psi_\infty(\tau) = \sum_{n=0}^{\infty} \varepsilon_n h_{n0}^2 \cos n p \tau. \quad (4.9-8)$$

Сравнивая это с известным уже результатом, а именно, что для

$$A + C \cos(2\pi f_0 t - \varphi), \quad (2.2-2)$$

функция корреляции равна

$$A^2 + \frac{C^2}{2} \cos 2\pi f_0 \tau, \quad (2.2-3)$$

и помня, что $\varepsilon_0=1$, а $\varepsilon_n=2$ для $n \geq 1$, получим

$$\begin{aligned} \text{Амплитуда постоянной составляющей тока } I &= h_{00}, \\ \text{Амплитуда } \frac{np}{2\pi} \text{-ой составляющей тока } I &= h_{n0}. \end{aligned} \quad (4.9-9)$$

Кстати, эти выражения для амплитуд почти сразу получаются при прямом методе решения, что будет показано в связи с уравнением (4.9—17).

Так как функция корреляции $\Psi_c(\tau)$ для сплошного участка $W_c(f)$ энергетического спектра тока I определяется как

$$\Psi_c(\tau) = \Psi(\tau) - \Psi_\infty(\tau), \quad (4.8-11)$$

то также получаем

$$\Psi_c(\tau) = \sum_{n=0}^{\infty} \sum_{k=1}^{\infty} \frac{1}{k!} \psi_\tau^k h_{nk}^2 \varepsilon_n \cos n p \tau. \quad (4.9-10)$$

Если это подставить в уравнение

$$W_c(f) = 4 \int_0^{\infty} \Psi_c(\tau) \cos 2\pi f\tau \, d\tau, \quad (4.9-11)$$

то получим

$$W_c(f) = \sum_{n=-\infty}^{\infty} \sum_{k=1}^{\infty} \frac{2\varepsilon_n}{k!} h_{nk}^2 \left[G_k\left(f - \frac{np}{2\pi}\right) + G_k\left(f + \frac{np}{2\pi}\right) \right], \quad (4.9-12)$$

где

$$G_k(f) = \int_0^{\infty} \psi_{\tau}^k \cos 2\pi f\tau \, d\tau \quad (4.9-13)$$

есть функция, исследованная в Приложении 4С и являющаяся четной функцией f . Двойной ряд (4.9—12) для W_c выглядит довольно сложным. Однако если представляет интерес только какой-то определенный участок частотного спектра, то часто достаточно учесть лишь небольшое число членов ряда.

Выше упоминалось, что прямой метод нахождения выходного энергетического спектра тесно связан с только что выведенными выражениями. Исследуем теперь эту связь.

Начнем со следующего вывода теории комбинационных тонов. Пусть к нелинейному устройству типа (4А—1) подведено напряжение

$$V = P_0 \cos x_0 + P_1 \cos x_1 + \dots + P_N \cos x_N; \\ x_k = p_k t, \quad k = 0, 1, \dots, N, \quad (4.9-14)$$

где p_k некрatны друг другу. Выходной ток равен

$$I = \sum_{m_0=0}^{\infty} \dots \sum_{m_N=0}^{\infty} \frac{1}{2} A_{m_0 \dots m_N} \varepsilon_{m_0} \\ \dots \varepsilon_{m_N} \cos m_0 x_0 \cos m_1 x_1 \dots \cos m_N x_N, \quad (4.9-15)$$

где $\varepsilon_0 = 1$, а при $m \geq 1$ $\varepsilon_m = 2$. Если произведение косинусов представить в виде суммы косинусов углов $m_0 x_0 \pm m_1 x_1 \dots \pm m_N x_N$, то видно, что коэффициент при типичном члене есть $A_{m_0 \dots m_N}$, за исключением случая, когда все m равны нулю. Тогда этот коэффициент равен $\frac{1}{2} A_{0 \dots 0}$. Поэтому

$$\frac{1}{2} A_{0 \dots 0} \text{— постоянная составляющая тока } I, \\ |A_{m_0 \dots m_N}| \text{— амплитуда составляющей частоты} \\ \frac{1}{2\pi} \left| m_0 p_0 \pm m_1 p_1 \pm \dots \pm m_N p_N \right|. \quad (4.9-16)$$

Для всех значений m

$$A_{m_0 \dots m_N} = \frac{iM}{\pi} \int_C F(iu) \prod_{r=0}^N J_{m_r}(P_r u) du, \quad (4.9-17)$$

$$M = m_0 + m_1 + \dots + m_N.$$

Заменим (4.9—14) на

$$V = P \cos pt + V_N, \quad (4.1-13)$$

полагая $P_0 = P$, $p_0 = p$ и представляя напряжение шумов V_N в виде суммы остальных членов. Так как при этом P_1, \dots, P_N будут весьма малыми, то способ Лапласа показывает, что в (4.9—17) можно положить

$$\prod_{r=1}^N J_0(P_r u) \approx \exp \left[-\frac{u^2}{4} (P_1^2 + \dots + P_N^2) \right] \approx e^{-\frac{\psi_0 u^2}{2}} \quad (4.9-18)$$

Здесь было использовано то обстоятельство, что ψ_0 есть средний квадрат напряжения V_N . Из этих уравнений следует:

$$\text{Постоянная составляющая } I = \frac{1}{2\pi} \int_C F(iu) J_0(Pu) e^{-\frac{\psi_0 u^2}{2}} du,$$

$$\text{Составляющая частоты } \frac{np}{2\pi} = \frac{i^n}{\pi} \int_C F(iu) J_n(Pu) e^{-\frac{\psi_0 u^2}{2}} du.$$

Эти результаты идентичны с (4.9—9).

Выведенные уравнения показывают, что h_{n0} должно быть связано с n -й гармоникой p . Подобным же образом можно доказать, что h_{nk} должно быть связано с комбинационными тонами создаваемыми n -й гармоникой частоты p и k элементарными синусоидальными составляющими напряжения V_N . Рассмотрим только комбинации типа $p_1 \pm p_2 \pm p_3$, взяв для примера $k=3$ и пренебрегая членами типа $3p_1$ и $2p_1 \pm p_2$. Комбинации первого из упомянутых выше типов значительно более многочисленны, а именно: число таких членов порядка N^3 , тогда как число членов последних двух типов соответственно N и N^2 .

Итак, примем $k=3$ и $m_1, m_2, m_3=1$, а $m_4, \dots, m_N=0$ в соответствии с комбинационными частотами вида $np \pm p_1 \pm p_2 \pm p_3$. Пользуясь приближениями, найдем

$$\begin{aligned} A_{n, 1, 1, 1, 0, 0, \dots, 0} &= \frac{i^{n+3}}{\pi} \frac{P_1 P_2 P_3}{8} \int_C F(iu) J_n(Pu) u^3 e^{-\frac{\psi_0 u^2}{2}} du = \\ &= \frac{P_1 P_2 P_3}{4} h_{n3}. \end{aligned}$$

Если рассматривается какая-либо другая комбинационная частота типа $np \pm p_{r_1} \pm p_{r_2} \pm p_{r_3}$, то получается подобное же выражение,

в котором $P_1 P_2 P_3$ заменено на $P_{r_1} P_{r_2} P_{r_3}$. Это может быть сделано для любого значения k . Полученный результат показывает, что h_{nk} и, следовательно, также (n, k) -ые члены в двойных рядах (4.9—10) и (4.9—12) для $\Psi_c(\tau)$ и $W_c(f)$ должны быть связаны с комбинационными частотами порядка (n, k) , где n относится к сигналу, а k — к составляющим шумов.

Теперь можно сформулировать теорему относительно полной мощности, связанной с комбинационными частотами данного порядка. Для выбранного нелинейного устройства (т. е. когда $F(iu)$ задано) полная мощность, которая рассеивается всеми комбинационными частотами порядка (n, k) , если ток I протекает через сопротивление 1 ом, равняется

$$\Psi_{nk}(0) = \frac{\varepsilon_n [\psi(0)]^k}{k!} h_{nk}^2 = \frac{\varepsilon_n [V_N^2]^k h_{nk}^2}{k!}. \quad (4.9-19)$$

Существенной особенностью этого выражения является то, что оно зависит только от эффективного значения напряжения V_N и от функции $F(iu)$ и вовсе не зависит от спектрального распределения шумов на входе.

Доказательство (4.9—19) основано на соотношении

$$\Psi_{nk}(0) = \int_0^{\infty} W_{nk}(f) df$$

между полной мощностью всех комбинационных частот порядка (n, k) и соответствующей функцией корреляции, найденной из (4.9—7).

Эта теорема была применена Миддлтоном для доказательства того, что если входное напряжение ограничено относительно узкой полосой частот, так что выходной спектр состоит из ряда полос, то мощность в каждой полосе частот зависит только от V_N^2 , но не от спектра V_N .

4.10. НЕКОТОРЫЕ РЕЗУЛЬТАТЫ, ПОЛУЧЕННЫЕ МЕТОДОМ ФУНКЦИИ КОРРЕЛЯЦИИ

В этом разделе приводится ряд выводов, которые могут быть получены из теории, изложенной в разделах, следующих за 4.6.

Если напряжение на входе квадратичного детектора, описываемого уравнением

$$I = \alpha V^2, \quad (4.1-1)$$

состоит только из одних шумов, так что $V = V_N$, то функция корреляции для тока I равна

$$\Psi(\tau) = \alpha^2 (\psi_0^2 + 2\psi_\tau^2), \quad (4.10-1)$$

где ψ_τ есть функция корреляции для V_N . Сравним (4.10—1) с уравнением (3.9—7). Если V взять в общем виде, то

$$\begin{aligned} \Psi(\tau) &= \overline{I(t)I(t+\tau)} = \alpha^2 \overline{V^2(t)V^2(t+\tau)} = \\ &= \alpha^2 \times \left[\begin{array}{l} \text{коэфф. при } \frac{(iu)^2}{2!} \frac{(iv)^2}{2!} \text{ в разложении в степенной ряд} \\ \text{характеристической функции для } V(t), V(t+\tau) \end{array} \right], \end{aligned} \quad (4.10-2)$$

где было использовано известное свойство характеристической функции. Выражение для х. ф., обозначенной $g(u, v, \tau)$, дается уравнением (4.8—4). Например, если напряжение V состоит из синусоидального напряжения и шумов (4.1—13), то характеристическая функция находится из (4.8—3) и (4.9—2).

Следовательно:

$$\begin{aligned} \Psi(\tau) &= \left\{ \begin{array}{l} \text{коэфф. при } \frac{u^2 v^2}{4} \text{ в разложении выражения} \\ \alpha^2 J_0(PV\sqrt{u^2+v^2+2uv \cos p\tau}) \exp \left[-\frac{\psi_0}{2}(u^2+v^2) - \psi_\tau uv \right] \end{array} \right\} = \\ &= \alpha^2 \left[\left(\frac{P}{2} + \psi_0 \right)^2 + \frac{P^4}{8} \cos 2p\tau + 2P^2 \psi_\tau \cos p\tau + 2\psi_\tau^2 \right]. \end{aligned} \quad (4.10-3)$$

Первые два члена дают постоянную составляющую и вторую гармонику. Последними двумя членами можно воспользоваться для вычисления сплошного участка энергетического спектра по уравнению (4.5—13).

В качестве примера приложения теории раздела 4.9 рассмотрим случай, когда синусоидальное напряжение вместе с напряжениями шумов (4.1—13) подается на вход детектора со следующей характеристикой:

$$\begin{aligned} I=0, \quad V < 0, \\ I=V^\nu, \quad V > 0. \end{aligned} \quad (4.10-4)$$

В таблице Приложения 4А находим

$$F(iu) = \Gamma(\nu+1) (iu)^{-\nu-1}.$$

Путь интегрирования C проходит вдоль вещественной оси от $-\infty$ до $+\infty$, отступая книзу у начала координат. Тогда интеграл (4.9—6) для h_{nk} приобретет вид

$$\begin{aligned} h_{nk} &= \frac{i^{n+k-\nu-1}}{2\pi} \Gamma(\nu+1) \int_C u^{k-\nu-1} J_n(Pu) e^{-\frac{\psi_0}{2}u^2} du = \\ &= \frac{\left(\frac{\psi_0}{2} \right)^{\frac{\nu-k}{2}} x^{\frac{n}{2}} \Gamma(\nu+1)}{2\Gamma\left(\frac{2-k-n+\nu}{2} \right) n!} {}_1F_1 \left(\frac{k+n-\nu}{2}; n+1; -x \right), \end{aligned} \quad (4.10-5)$$

где $x = \frac{P^2}{2\psi_0}$, интегрирование выполнено путем разложения $J_n(Pu)$ по степеням u и использования соотношения

$$\int_C e^{-au^2} u^{2\lambda-1} du = ie^{-\lambda i\pi} a^{-\lambda} \sin \lambda\pi \Gamma(\lambda) = \\ = \frac{a^{-\lambda}}{2} \left(1 - e^{-2\lambda i\pi}\right) \Gamma(\lambda) = \frac{i\pi e^{-\lambda i\pi}}{a^\lambda \Gamma(1-\lambda)}. \quad (4.10-6)$$

Здесь подразумевается, что $\arg u=0$ в положительной части C .

Согласно (4.9—9), постоянная составляющая I дается уравнением

$$h_{00} = \frac{\Gamma(1+\nu)}{2\Gamma\left(1+\frac{\nu}{2}\right)} \left(\frac{\psi_0}{2}\right)^{\nu/2} {}_1F_1\left(-\frac{\nu}{2}; 1; -x\right) \quad (4.10-7)$$

которое сводится к (4.2—3), когда $\nu=1$ для линейного детектора (за исключением коэффициента α).

Если входное напряжение (синусоидальное напряжение и шумы) ограничено относительно узкой полосой частот, а нас интересует низкочастотное выходное напряжение, то следует рассмотреть различные комбинационные тона, соответствующие комбинационным частотам порядка (0,0), (0,2), (0,4),..... (1,1), (1,3),....., (2,0), (2,2)... и т. д., где типичные частоты имеют порядок (n,k) . Члены порядка (0,0) и (2,0) дают постоянную составляющую и вторую гармонику и, следовательно, не учитываются при вычислении $W_c(f)$. Из оставшихся членов наибольшее значение в рядах (4.9—12) и (4.9—10) для $W_c(f)$ и $\Psi_c(\tau)$ имеют либо (0,2), либо (1,1). Другие члены имеют все меньшее значение по мере возрастания n и k . Тогда низкочастотная часть сплошного участка выходного энергетического спектра по (4.9—12) равна

$$W_c(f) = \frac{4}{2!} h_{02}^2 G_2(f) + \frac{4}{4!} h_{04}^2 G_4(f) + \dots + \\ + \frac{4}{1!} h_{11}^2 [G_1(f-f_0) + G_1(f+f_0)] + \frac{4}{3!} h_{13}^2 [G_3(f-f_0) + \\ + G_3(f+f_0)] + \frac{4}{2!} h_{22}^2 [G_2(f-2f_0) + G_2(f+2f_0)] + \dots \quad (4.10-8)$$

Из табл. 2 в Приложении 4С можно определить низкочастотные части функций G . Следует помнить, что $G_m(x)$ есть четная функция x и что $0 < f \ll f_0$.

В качестве примера пусть напряжение входных шумов V_N имеет такие же $\omega(f)$ и $\psi(\tau)$, что и в случае фильтра a (фильтр

с вероятностной характеристикой) в Приложении 4С, так что ¹⁾

$$w(f) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(f-f_0)^2/2\sigma^2}$$

Допустим, что частота синусоидального сигнала лежит в середине полосы частот, т. е. $p=2\pi f_0$. Поэтому из (4.10—8) в случае распределения входных шумов, следующего нормальному закону, для участка низких частот найдем:

$$\begin{aligned} W_c(f) = & \frac{1}{4\sigma\sqrt{\pi}} h_{02}^2 \psi_0^2 e^{-f^2/4\sigma^2} + \frac{1}{64\sigma\sqrt{2\pi}} h_{04}^2 \psi_0^4 e^{-f^2/8\sigma^2} + \\ & + \frac{2}{\sigma\sqrt{2\pi}} h_{11}^2 \psi_0 e^{-f^2/2\sigma^2} + \frac{1}{4\sigma\sqrt{6\pi}} h_{13}^2 \psi_0^3 e^{-f^2/6\sigma^2} + \\ & + \frac{1}{4\sigma\sqrt{\pi}} h_{22}^2 \psi_0^2 e^{-f^2/4\sigma^2} + \dots \end{aligned} \quad (4.10-9)$$

Хотя здесь рассматривался детектор с характеристикой степени ν , уравнение (4.10—9) дает низкочастотный участок спектра $W_c(f)$, соответствующий шумам, подчиняющимся нормальному закону, для любого нелинейного устройства, если только в уравнение подставлены надлежащие h_{nk} .

Частный случай уравнения (4.10—9) возникает тогда, когда на линейный детектор подаются одни шумы. Низкочастотный участок выходного энергетического спектра

$$\begin{aligned} W_c(f) = & \frac{\psi_0}{\pi} \sum_{m=1}^{\infty} \frac{\left(-\frac{1}{2}\right)_m \left(-\frac{1}{2}\right)_m}{m! m!} \frac{1}{\sigma\sqrt{4m\pi}} e^{-f^2/4m\sigma^2} = \\ = & \frac{\psi_0 \pi^{-3/2}}{2\sigma} \left[\frac{1}{4} e^{-f^2/4\sigma^2} + \frac{1}{64\sqrt{2}} e^{-f^2/8\sigma^2} + \right. \\ & \left. + \frac{1}{256\sqrt{3}} e^{-f^2/12\sigma^2} + \dots \right] \end{aligned} \quad (4.10-10)$$

где были применены уравнение (4.7—6) и табл. 2 из Приложения 4С. Функция корреляции для

$$V_s = P \cos pt + Q \cos qt,$$

где p и q — некрáтные величины, есть

$$J_0(P\sqrt{u^2 + v^2 + 2uv \cos p\tau}) J_0(Q\sqrt{u^2 + v^2 + 2uv \cos q\tau}).$$

¹⁾ См. примечание на стр. 167. (Прим. ред.)

Из уравнений (4.9—16) и (4.9—17) немедленно следует, что

$$h_{000} = \frac{1}{2\pi} \int_C F(iu) J_0(Pu) J_0(Qu) e^{-(u^2/2)\psi_0} du \quad (4.10-11)$$

есть постоянная составляющая тока I , если приложенное напряжение равно

$$P \cos pt + Q \cos qt + V_N. \quad (4.1-4)$$

Рагадзини получил приближенное выражение для выходного энергетического спектра, когда напряжение

$$\begin{aligned} V &= V_s + V_N, \\ V_s &= Q(1 + r \cos pt) \cos qt \end{aligned} \quad (4.10-12)$$

подведено к линейному детектору¹⁾. В наших обозначениях его выражение для сплошного участка энергетического спектра (для области низких частот)

$$W_c(f) = \frac{1}{\pi^2 \alpha^2 (Q^2 + 2\psi_0)} \times \left[\begin{array}{l} W_c(f) \text{ по уравнению (4.5-16)} \\ \text{для квадратичного детектора} \end{array} \right] \quad (4.10-13)$$

В знаменатель подставлено α^2 , чтобы сократить α^2 в выражении (4.5—16). Для линейного детектора было положено

$$I = \begin{cases} 0, & V < 0, \\ V, & 0 < V, \end{cases} \quad (4.10-14)$$

и индекс модуляции k в (4.5—16) заменен на r .

Формула Рагадзини совершенно точна, если индекс модуляции r мал, особенно когда $y = Q^2/(2\psi_0)$ велико. Чтобы доказать это, положим в (4.10—13) $r=0$; тогда

$$W_c(f) = \frac{1}{\pi^2 (Q^2 + 2\psi_0)} \left[\begin{array}{l} Q^2 w(f_q - f) + Q^2 w(f_q + f) + \\ + \int_{-\infty}^{+\infty} w(x) w(f - x) dx \end{array} \right] \quad (4.10-15)$$

где $f_q = q/2\pi$. Это выражение нужно сравнить с низкочастотной частью $W_c(f)$ из уравнения (4.10—8) для частного случая линейного детектора, на входе которого действуют синусоидальное напряжение и шумы. Основные члены в (4.10—8) дают

$$\begin{aligned} W_c(f) &= h_{11}^2 [\dot{w}(f_q - f) + \dot{w}(f_q + f)] + \\ &+ h_{02}^2 \frac{1}{4} \int_{-\infty}^{+\infty} w(x) w(f - x) dx. \end{aligned} \quad (4.10-16)$$

¹⁾См. уравнение (12) в статье «Действие флуктуационных напряжений на линейный детектор», PIRE, 30, 277—288, июнь 1942.

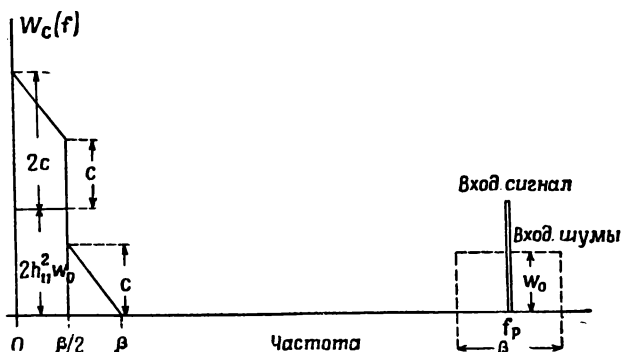
Значения h , соответствующие линейному детектору, получаются, если положить в (4.10—5) $\nu=1$, помня, что Q теперь играет роль P :

$$h_{11} = \frac{1}{2} \left(\frac{y}{\pi} \right)^{1/2} {}_1F_1 \left(\frac{1}{2}; 2; -y \right)$$

$$h_{02} = (2\pi\psi_0)^{-1/2} {}_1F_1 \left(\frac{1}{2}; 1; -y \right) \quad (4.10-17)$$

$$y = \frac{Q^2}{2\psi_0}$$

Кстати, первое приближение для выходного эффекта линейного детектора, даваемое уравнением (4.10—16), интересно само по себе. На фиг. 9 показан низкочастотный участок $W_c(f)$, вычислен-



Фиг. 9. Низкочастотный участок спектра на выходе линейного детектора.

$$c = h_{02}^2 \frac{\beta \omega_0^2}{4} = (Ph_{02})^2 \left(\frac{\beta \omega_0^2}{4P^2} \right).$$

Постоянная составляющая на выходе $= Ph_{11} + \beta \omega_0 h_{02}$.

ный по уравнению (4.10—16) для случая, когда шумы на входе равномерно распределены в узкой полосе частот шириной β , а f_p — средняя частота этой полосы частот. h_{11} и h_{02} можно определить из кривых фиг. 10. На фиг. 9 и графиках фиг. 10 P и x заменяют Q и y , входящие в (4.10—17), чтобы сохранить те же обозначения, что и на фиг. 8 для квадратичного детектора.

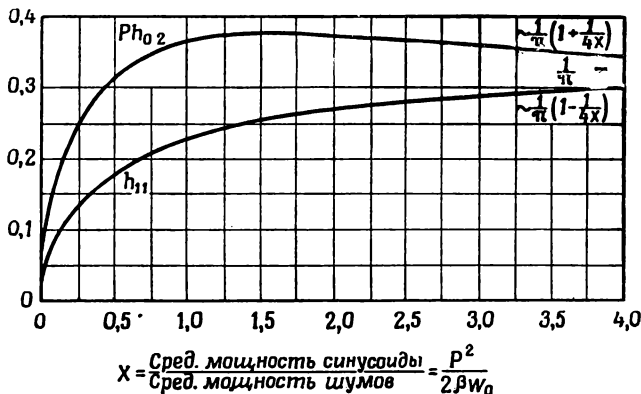
Для сравнения полезны следующие значения:

когда $x = 0$	когда x велико	
$h_{11} = 0$	$h_{11} = \frac{1}{\pi}$	
$h_{02} = (2\pi\psi_0)^{-1/2}$	$h_{02} = \frac{1}{\pi Q}$	(4.10—18)

Значения для больших x найдены из асимптотического разложения (4В—3) в Приложении 4В.

Первое сравнение между (4.10—15) и (4.10—16) произведем, полагая $Q \rightarrow \infty$. Тогда оба выражения приводятся к

$$W_c(f) = \frac{1}{\pi^2} \left[\omega(f_q - f) + \omega(f_q + f) \right] \quad (4.10—19)$$



Фиг. 10. Коэффициенты для расчета выходного эффекта линейного детектора согласно фиг. 9.

$$Ph_{02} = \sqrt{\frac{x}{\pi}} {}_1F_1\left(\frac{1}{2}; 1; -x\right),$$

$$h_{11} = \frac{1}{2} \sqrt{\frac{x}{\pi}} {}_1F_1\left(\frac{1}{2}; 2; -x\right).$$

что показывает, что в этом случае имеет место полное совпадение. Пусть теперь $Q=0$. Тогда оба уравнения дают

$$W_c(f) = \frac{1}{A 2\pi\psi_0} \int_{-\infty}^{+\infty} \omega(x)\omega(f-x) dx,$$

где $A=\pi$ для формулы Рагадзини и $A=4$ для (4.10—16). Поэтому результаты вычислений все еще вполне хорошо согласуются. Предельное значение (4.10—16) может быть также получено из (4.7—8).

Даже если индекс модуляции r и не совсем мал, то можно показать, что, когда $Q \rightarrow \infty$, $W_c(f)$ все еще достигает значений, определяемых уравнением (4.10—19). Формула Рагадзини дает несколько больший результат, так как она включает дополнительные члены, показанные в (4.5—17), которые содержат $k^2/4$, но эта разница не кажется серьезной. Если $Q^2 + 2\psi_0$ в знаменателе (4.10—13) заменить на $Q^2 + \frac{1}{2}Q^2k^2 + 2\psi_0$, то согласие между формулами улучшается.

Приложение 4А

Таблица нелинейных устройств, описываемых интегралами

Многие нелинейные устройства могут быть описаны интегралами вида

$$I = \frac{1}{2\pi} \int_C F(iu) e^{iVu} du, \quad (4A-1)$$

где функция $F(iu)$ и путь интегрирования C выбираются в соответствии с нелинейным устройством. В таблице даны примеры подобных устройств. Некоторые важные случаи не могут быть представлены в такой форме. Примером является ограничитель с характеристикой

$$\begin{aligned} I &= -\alpha D, & V < -D, \\ I &= \alpha V, & -D < V < D, \\ I &= \alpha D, & D < V, \end{aligned} \quad (4A-2)$$

для которого можно написать

$$\begin{aligned} I &= \frac{2\alpha}{\pi} \int_0^{\infty} \sin Vu \sin Du \frac{du}{u^2} = \\ &= -\alpha D + \frac{2\alpha}{2\pi i} \int_C e^{iVu} \sin Du \frac{du}{u^2}, \end{aligned} \quad (4A-3)$$

где C простирается от $-\infty$ до $+\infty$ и отступает книзу у начала координат. Эта форма интеграла отличается от рассматриваемой в главе IV. Однако, кажется, не представит трудностей распространение теории на частный случай ограничителя.

НЕЛИНЕЙНЫЕ УСТРОЙСТВА, ОПИСЫВАЕМЫЕ ИНТЕГРАЛАМИ

$$I = \frac{1}{2\pi} \int_C F(iu) e^{iVu} du$$

I	$F(iu)$	C	Тип устройства
$I = \alpha V^n,$ n — целое	$\frac{\alpha n!}{(iu)^{n+1}}$	Положительный контур вокруг $u=0$	Устройство с характеристикой n -й степени
$I = \alpha(V-B)^n,$ n — целое	$\frac{\alpha n!}{(iu)^{n+1}} e^{-iVB}$	Положительный контур вокруг $u=0$	Устройство с характеристикой n -й степени со смещением

Продолжение

I	$F(iu)$	C	Тип устройства
$I=0, V<0$ $I=aV, 0<V$	$\frac{\alpha}{(iu)^2} = -\frac{\alpha}{u^2}$	Вещественная u ось от $-\infty$ до $+\infty$ с отступлением книзу у $u=0$	Линейный детектор; точка среза при $V=0$
$I=0, V<B$ $I=a(V-B)^\nu$, $V>B$ ν —любое положительн. число	$\frac{\alpha\Gamma(\nu+1)}{(iu)^{\nu+1}} e^{-iuB}$	»	Детектор с характеристикой ν -й степени со смещением
$I=0, V<0$ $I=aV, 0<V<D$ $I=aD, D<V$	$\frac{\alpha(1-e^{-iuD})}{(iu)^2}$	»	Линейный детектор плюс ограничитель
$I=0, V<0$ $I=\varphi(V), V>0$	$F(p) = \int_0^\infty e^{-pt} \varphi(t) dt$		

Приложение 4В

Функция ${}_1F_1(a; c; x)$

В вопросах, связанных с прохождением через нелинейные устройства синусоидальных колебаний вместе с шумами, появляется гипергеометрическая функция

$${}_1F_1(a; c; z) = 1 + \frac{az}{c1!} + \frac{a(a+1)}{c(c+1)} \frac{z^2}{2!} + \dots \quad (4B-1)$$

Здесь будут изложены некоторые ее свойства, используемые в главе IV. Кривые ${}_1F_1(a; c; z)$ даны для $a=-4; -3,5\dots; 3,5; 4,0$ и $c=-1,5; -0,5; +0,5; 1; 1,5; 2; 3; 4$ в книге Янке и Эмде «Таблицы функций»¹⁾. Приведен также список свойств этой функции и другие ссылки.

Если c не является отрицательным целым числом или нулем, то

$${}_1F_1(a; c; z) = e^z {}_1F_1(c-a; c; -z). \quad (4B-2)$$

¹⁾ См. Янке и Эмде, «Таблицы функций», Гостехиздат, 373, 1948.

Когда $\operatorname{Re}(z) > 0$, то справедливы асимптотические разложения

$$\begin{aligned}
 {}_1F_1(a; c; z) &\approx \frac{\Gamma(c) e^z}{\Gamma(a) z^{c-a}} \left[1 + \frac{(1-a)(c-a)}{1! z} + \right. \\
 &+ \left. \frac{(1-a)(2-a)(c-a)(c-a+1)}{2! z^2} + \dots \right], \\
 {}_1F_1(a; c; -z) &\approx \frac{\Gamma(c)}{\Gamma(c-a) z^a} \left[1 + \frac{a(1+a-c)}{1! z} + \right. \\
 &+ \left. \frac{a(a+1)(1+a-c)(2+a-c)}{2! z^2} + \dots \right].
 \end{aligned} \quad (4B-3)$$

Многие из гипергеометрических функций, с которыми приходится сталкиваться, могут быть выражены через функции Бесселя первого рода мнимого аргумента. Эта связь может быть установлена при помощи соотношения¹⁾

$${}_1F_1\left(\nu + \frac{1}{2}; 2\nu + 1; z\right) = 2^{2\nu} \Gamma(\nu + 1) z^{-\nu} e^{z/2} I_\nu\left(\frac{z}{2}\right) \quad (4B-4)$$

вместе с рекуррентными соотношениями

	F_{a+}	F_{a-}	F_{c+}	F_{c-}	F
1.	a	$(a-c)$			$c-2a-z$
2.	ac		$(c-a)z$		$-c(a+z)$
3.	a			$1-c$	$c-a-1$
4.		$-c$	$-z$		c
5.		$a-c$		$c-1$	$1-a-z$
6.			$(c-a)z$	$c(c-1)$	$c(1-c-z)$

Например, первое рекуррентное соотношение получается из первой строчки в следующем виде:

$$\begin{aligned}
 aF(a+1; c; z) + (a-c)F(a-1; c; z) + \\
 + (c-2a-z)F(a; c; z) = 0.
 \end{aligned} \quad (4B-5)$$

Эти шесть соотношений между смежными ${}_1F_1$ функциями аналогичны 15 соотношениям, полученным для смежных ${}_2F_1$ гипергеометрических функций, и могут быть выведены из последних при помощи соотношения

$${}_1F_1(a; c; z) = \lim_{b \rightarrow \infty} {}_2F_1\left(a; b; c; \frac{z}{b}\right). \quad (4B-6)$$

Рекуррентное соотношение, вводящее две функции ${}_1F_1$ типа (4B-4), можно получить, заменяя в соотношении, даваемом четвер-

¹⁾ Г. Н. Ватсон, «Теория бесселевых функций», ГИИЛ, 1949.

той строчкой таблицы, a на $a+1$, а затем устранив ${}_1F_1(a+1; c; z)$ из этого соотношения, а также из полученного с помощью третьей строчки таблицы. Это дает

$${}_1F_1(a; c; z) = {}_1F_1(a; c-1; z) + \frac{za}{c(1-c)} F(a+1; c+1; z). \quad (4B-7)$$

Полагая ν равным нулю и единице в (4B-4) и $a = \frac{1}{2}$, $c=2$ в (4B-7), получим

$$\begin{aligned} {}_1F_1\left(\frac{1}{2}; 1; z\right) &= e^{z/2} I_0\left(\frac{z}{2}\right), \\ {}_1F_1\left(\frac{3}{2}; 3; z\right) &= 4z^{-1} e^{z/2} I_1\left(\frac{z}{2}\right), \\ {}_1F_1\left(\frac{1}{2}; 2; z\right) &= e^{z/2} \left[I_0\left(\frac{z}{2}\right) - I_1\left(\frac{z}{2}\right) \right]. \end{aligned} \quad (4B-8)$$

Соотношения в приведенной таблице позволяют найти выражения для ${}_1F_1\left(n + \frac{1}{2}; m; z\right)$, где n и m — целые числа. В частности, воспользовавшись (4B-2), получим

$${}_1F_1\left(-\frac{1}{2}; 1; -z\right) = e^{-z/2} \left[(1+z) I_0\left(\frac{z}{2}\right) + z I_1\left(\frac{z}{2}\right) \right]. \quad (4B-9)$$

Приложение 4С'

Энергетический спектр, соответствующий ψ_τ^n

Весьма часто встречается интеграл

$$G_n(f) = \int_0^\infty [\psi(\tau)]^n \cos 2\pi f\tau \, d\tau, \quad (4C-1)$$

где $\psi(\tau)$ есть функция корреляции, соответствующая энергетическому спектру $w(f)$. Из основного соотношения между $w(f)$ и $\psi(\tau)$ согласно уравнению (2.1-5)

$$G_1(f) = \frac{w(f)}{4}. \quad (4C-2)$$

Выражение для спектра произведения двух функций позволяет представить $G_n(f)$ через $w(f)$. Будем пользоваться следующей формой этого выражения. Пусть $F_r(f)$ будет спектром функции $\varphi_r(\tau)$, так что

$$\dot{\varphi}_r(\tau) = \int_{-\infty}^{+\infty} F_r(f) e^{2\pi i f\tau} \, df, \quad r = 1, 2,$$

$$F_r(f) = \int_{-\infty}^{+\infty} \varphi_r(\tau) e^{-2\pi i f \tau} d\tau.$$

Тогда

$$\int_{-\infty}^{+\infty} \varphi_1(\tau) \varphi_2(\tau) e^{-2\pi i f \tau} d\tau = \int_{-\infty}^{+\infty} F_1(x) F_2(f-x) dx, \quad (4C-3)$$

т. е. спектр произведения $\varphi_1(\tau)\varphi_2(\tau)$ есть интеграл, написанный справа.

Если $\varphi_1(\tau)$ и $\varphi_2(\tau)$ — вещественные четные функции τ , то (4C—3) можно написать в виде

$$\int_0^{\infty} \varphi_1(\tau) \varphi_2(\tau) \cos 2\pi f \tau d\tau = \frac{1}{2} \int_{-\infty}^{+\infty} F_1(x) F_2(f-x) dx. \quad (4C-4)$$

Чтобы получить $G_2(f)$, положим $\varphi_1(\tau)$ и $\varphi_2(\tau)$ равными $\psi(\tau)$. Тогда можно воспользоваться (4C—4), так как $\psi(\tau)$ есть четная вещественная функция τ . Если $\varphi_r(\tau)$ есть четная вещественная функция τ , то из формулы интеграла Фурье для $F_r(f)$ следует, что $F_r(f)$ должно быть четной вещественной функцией f . Поэтому положим

$$2F_r(f) = \omega(f), \quad r = 1, 2$$

и определим спектр $\omega(f)$ для отрицательных f как

$$\omega(-f) = \omega(f). \quad (4C-5)$$

Тогда уравнение (4C—4) дает

$$\begin{aligned} G_2(f) &= \frac{1}{8} \int_{-\infty}^{+\infty} \omega(x)\omega(f-x) dx = \\ &= \frac{1}{8} \int_0^{\infty} \omega(x)\omega(f-x) dx + \frac{1}{8} \int_0^{\infty} \omega(x)\omega(f+x) dx, \end{aligned} \quad (4C-6)$$

где во втором уравнении появляются только положительные значения аргумента $\omega(f)$.

Чтобы определить $G_3(f)$, положим $\varphi_1(\tau) = \psi(\tau)$, $2F_1(f) = \omega(f)$ и $\varphi_2(\tau) = \psi^2(\tau)$. Тогда

$$F_2(f) = 2 \int_0^{\infty} \varphi_2(\tau) \cos 2\pi f \tau d\tau = 2G_2(f)$$

и из (4С—4) получим

$$G_3(f) = \frac{1}{2} \int_{-\infty}^{+\infty} w(x) G_2(f-x) dx =$$

$$= \frac{1}{16} \int_{-\infty}^{+\infty} w(x) dx \int_{-\infty}^{+\infty} w(y) w(f-y) dy. \quad (4C-7)$$

Уравнение (4С—7) наводит на мысль, что выражение для $G_2(f)$ можно написать в виде

$$G_2(f) = \frac{1}{2} \int_{-\infty}^{+\infty} w(x) G_1(f-x) dx. \quad (4C-8)$$

То, что это справедливо, видно из (4С—2) и (4С—6). В результате оказывается, что уравнение

$$G_n(f) = \frac{1}{2} \int_{-\infty}^{+\infty} w(f-x) G_{n-1}(x) dx \quad (4C-9)$$

может быть использовано для поэтапного вычисления $G_n(f)$.

Рассмотрим теперь $G_n(f)$ для случая полосовых фильтров с относительно узкой полосой пропускания. В качестве примера возьмем фильтры, характеристики которых дают следующие значения $w(f)$ и $\psi(\tau)$.

Таблица 1

Фильтр	$w(f)$ для $f > 0$	$\psi(\tau)$
<i>a</i>	$\frac{\psi_0}{\sigma \sqrt{2\pi}} e^{-(f-f_0)^2/2\sigma^2}$	$\psi_0 e^{-2(\pi\sigma\tau)^2} \cos 2\pi f_0 \tau$
<i>b</i>	$\frac{\psi_0 \alpha}{\pi} \frac{1}{\alpha^2 + (f-f_0)^2}$ $w(f) = w_0 = \psi_0/\beta$ для	$\psi_0 e^{-2\pi\alpha \tau } \cos 2\pi f_0 \tau$
<i>c</i>	$\left(f_0 - \frac{\beta}{2}\right) < f < \left(f_0 + \frac{\beta}{2}\right)$ $w(f) = 0$ для остальных частот	$\psi_0 \frac{\sin \pi\beta\tau}{\pi\beta\tau} \cos 2\pi f_0 \tau$

Эти фильтры будем называть соответственно фильтр *a*, фильтр *b* и фильтр *c*. Во всех фильтрах f_0 — средняя частота полосы

пропускания. Постоянные выбраны так, что когда приложено напряжение, занимающее широкую полосу частот, все фильтры пропускают одну и ту же среднюю мощность

$$\psi_0 = \int_0^{\infty} \omega(f) df = \text{средний квадрат } V(t) \text{ или } I(t),$$

причем предполагается, что $f_0 \gg \sigma$, $f_0 \gg \alpha$, $f_0 \gg \beta$, так что полосы пропускания сравнительно узки.

Выражения для $G_n(f)$, соответствующие нескольким значениям n , приведены в табл. 2. Когда $n=1$, $G_1(f)$ просто равно $\omega(f)/4$. $G_2(f)$ получается, полагая $n=2$ в (4C—1) для $G_n(f)$, возводя в квадрат $\psi(\tau)$, взятое из табл. 1, и пользуясь соотношением

$$\cos^2 2\pi f_0 \tau = \frac{1}{2} + \frac{1}{2} \cos 4\pi f_0 \tau.$$

Выражение для $G_2(f)$ в табл. 2, относящееся к фильтру c , является точным. Выражения для случая фильтров a и b дают хорошее приближение вблизи $f=0$ и $f=2f_0$, когда $G_2(f)$ велико. Однако они не являются точными, так как отброшены члены, содержащие $f+2f_0$. Как следует из таблицы, все функции G_2 в рассматриваемых трех случаях ведут себя одинаково. Каждая функция имеет пик, симметричный вокруг $2f_0$, ширина которого в два раза больше, чем ширина пика в спектре $\omega(f)$; затем между 0 и $2f_0$ функция равна почти нулю, а при нуле создает опять пик, высота которого в два раза превышает пик при $2f_0$.

$G_3(f)$ получается путем возведения в куб взятых из табл. 1 значений $\psi(\tau)$, используя соотношение

$$\cos^3 2\pi f_0 \tau = \frac{3}{4} \cos 2\pi f_0 \tau + \frac{1}{4} \cos 6\pi f_0 \tau.$$

Рассматривая, каким образом в (4C—1) косинусоидальные члены комбинируются с $\cos 2\pi f\tau$, приходим к заключению, что для полосовых фильтров с относительно узкой полосой $G_3(f)$ имеет пики при f_0 и $3f_0$, причем первый пик в три раза больше второго. Выражения для $G_3(f)$ и $G_4(f)$ являются приближенными в том же самом смысле, что и выражения для $G_2(f)$. Можно заметить, что коэффициенты внутри скобок для фильтров a и b есть биномиальные коэффициенты для рассматриваемых значений n . Поэтому для $n=2$ они равны 2 и 1, для $n=3$ они равны 3 и 1 и для $n=4$ они равны 6, 4 и 1.

Функции $G_n(f)$ более высоких порядков для фильтров a и b могут быть вычислены подобным же образом. Встречающиеся интегралы равны

$$\int_0^{\infty} e^{-2n(\pi\sigma\tau)^2} \cos 2\pi f\tau d\tau = \frac{e^{-f^2/2n\sigma^2}}{2\sigma\sqrt{2n\pi}}$$

$G_n(f)$	Фильтр а	Фильтр б
$G_1(f)$	$\frac{\psi_0}{4\sigma\sqrt{2\pi}} e^{-(f-f_0)^2/2\sigma^2}$	$\frac{\alpha\psi_0}{4\pi} \frac{1}{\alpha^2 + (f-f_0)^2}$
$G_2(f)$	$\frac{\psi_0^2}{8\sigma\sqrt{4\pi}} [2e^{-f^2/4\sigma^2} + e^{-(f-2f_0)^2/4\sigma^2}]$	$\frac{2\alpha\psi_0^2}{8\pi} \left[\frac{2}{4\alpha^2 + f^2} + \frac{1}{4\alpha^2 + (f-2f_0)^2} \right]$
$G_3(f)$	$\frac{\psi_0^3}{16\sigma\sqrt{6\pi}} [3e^{-(f-f_0)^2/6\sigma^2} + e^{-(f-3f_0)^2/6\sigma^2}]$	$\frac{3\alpha\psi_0^3}{16\pi} \left[\frac{3}{9\alpha^2 + (f-f_0)^2} + \frac{1}{9\alpha^2 + (f-3f_0)^2} \right]$
$G_4(f)$	$\frac{\psi_0^4}{32\sigma\sqrt{8\pi}} [6e^{-f^2/8\sigma^2} + 4e^{-(f-2f_0)^2/8\sigma^2} + e^{-(f-4f_0)^2/8\sigma^2}]$	$\frac{4\alpha\psi_0^4}{32\pi} \left[\frac{6}{16\alpha^2 + f^2} + \frac{4}{16\alpha^2 + (f-2f_0)^2} + \frac{1}{16\alpha^2 + (f-4f_0)^2} \right]$
$G_n(f)$ n — нечетное f — мала	0	0
$G_n(f)$ n — четное f — мала	$\frac{\psi_0^n n!}{\left(\frac{n}{2}\right)! \left(\frac{n}{2}\right)!} \frac{e^{-f^2/2n\sigma^2}}{2^{n+1}\sigma\sqrt{2n\pi}}$	$\frac{\psi_0^n n!}{\left(\frac{n}{2}\right)! \left(\frac{n}{2}\right)! 2^n} \frac{1}{2^{n+1}\pi n \alpha} \frac{1}{1 + \left(\frac{f}{n\alpha}\right)^2}$
$G_n(f)$ n — четное n — велико f — мала	$\frac{1}{2\pi\sigma n} e^{-f^2/2n\sigma^2}$	$\frac{2}{\alpha(2\pi n)^{3/2}} \frac{1}{1 + \left(\frac{f}{n\alpha}\right)^2}$
Фильтр с $G_1(f)$	$\frac{\psi_0}{4\beta}$, когда $(f_0 - \frac{\beta}{2}) < f < (f_0 + \frac{\beta}{2})$ 0 при др. частотах	Фильтр с $G_2(f)$ $\frac{\psi_0^2}{4\beta} \left(1 - \frac{f}{\beta}\right)$, когда $0 < f < \beta$ $\frac{\psi_0^2}{8\beta^2} (f - 2f_0 + \beta)$, $(2f_0 - \beta) < f < 2f_0$ $\frac{\psi_0^2}{8\beta^2} (2f_0 + \beta - f)$, $2f_0 < f < (2f_0 + \beta)$

$$\int_0^{\infty} e^{-2\pi a \tau} \cos 2\pi f \tau d\tau = \frac{1}{2\pi} \frac{na}{n^2 a^2 + f^2}$$

Во многих примерах представляли интерес только значения $G_n(f)$ для f вблизи нуля, т. е. только тот пик, который образуется вблизи нуля. Очевидно, что $G_n(f)$ имеет подобный пик только тогда, когда n — четное; этот пик образуется благодаря наличию постоянного члена в разложении

$$\begin{aligned} \cos^{2k} x = & \frac{1}{2^{2k-1}} \left[\cos 2kx + 2k \cos 2(k-1)x + \right. \\ & + \frac{(2k)(2k-1)}{2!} \cos 2(k-2)x + \\ & \left. + \frac{(2k)!}{(k-1)!(k+1)!} \cos 2x + \frac{(2k)!}{k! k! 2} \right] \end{aligned}$$

ПРИМЕНЕНИЕ СТАТИСТИЧЕСКИХ МЕТОДОВ В РАДИОЛОКАЦИИ

Г л а в а I

ПРИНЦИП «ОБРАТНОЙ ВЕРОЯТНОСТИ» В ТЕОРИИ ПЕРЕДАЧИ СИГНАЛОВ¹⁾

Ф. ВУДВОРД и И. ДЭВИС

Теория получения данных подчеркивает значение относительных вероятностей возможных «сообщений», вычисленных перед и после приема сигнала. Эти последние вероятности объединяют все данные, несомые принятым сигналом, и могут быть вычислены из него при помощи «обратной вероятности». В принципе это вычисление может быть выполнено электронными устройствами в самом приемнике. Таким образом получаем определение оптимального устройства для извлечения всех сведений, остающихся в сигнале, на который воздействовали шумы. Указания на то, как эта теория может быть применена к задачам радиолокации и связи, даются в последнем параграфе работы.

1. «ОБРАТНАЯ ВЕРОЯТНОСТЬ» И ПЕРЕДАЧА СООБЩЕНИЙ

Любое пригодное к передаче сообщение является результатом выбора из множества возможных сообщений, которые могут быть либо дискретными, либо непрерывно переходить одно в другое. В теории передачи сообщений состояние знания перед приемом сигнала может быть описано указанием вероятности появления каждого сообщения. После приема сигнала можно надеяться, что из всей группы возможных сообщений будет выделено некоторое определенное сообщение. Другими словами, его априорная вероятность заменяется на апостериорную вероятность, равную единице.

Благодаря искажениям за счет помех принятый сигнал не всегда будет с полной достоверностью указывать переданное сообщение. Поэтому апостериорная вероятность не будет полностью сконцентрирована на одном сообщении, а будет распределена между несколь-

¹⁾Ph. M. Woodward and I. L. Davies, «Information Theory and Inverse Probability in Telecommunications», *Proc. of the IEE*, part III, 99, № 58, 37—44, March 1952.

кими сообщениями. Следовательно, в общем случае полная вероятность, равная единице, распределяется между сообщениями после приема сигнала иным способом, нежели перед приемом.

В этом заключается одна из существенных особенностей теории передачи электрических сигналов. Интуитивно мы полагаем, что увеличение вероятности истинного сообщения соответствует возрастанию количества данных. Прежде чем окажется возможным сформулировать это более точно, необходимо рассмотреть сами априорные и апостериорные распределения вероятностей, что немедленно приводит к принципу «обратной вероятности».

Наилучшим способом изложения этой теоремы является рассмотрение простого примера.

Допустим, что для передачи одного из двух сообщений, «да» и «нет», используется простая телеграфная система и что сообщения передаются двумя различными сигналами, отмечаемыми в приемнике зажиганием зеленой и красной ламп. Допустим далее, что в большом числе случаев, когда передается сообщение, помехи превращают определенную часть «красных» сигналов в «зеленые» и наоборот. Ради общности положим, что эти части различных, скажем, $\frac{2}{5}$ «зеленых» сигналов становятся «красными» и $\frac{1}{3}$ «красных» становятся «зелеными». Наконец, допустим, что среди передаваемых сообщений «да» и «нет» встречаются в отношении 5 : 3.

Тогда все имеющиеся сведения могут быть схематически записаны в виде табл. 1.

Таблица 1

Да	Да	Да	Да	Да	Нет	Нет	Нет
Зеленый	Зеленый	Зеленый	Красный	Красный	Красный	Красный	Зеленый

Эта таблица представляет просто перечисление всех равно-возможных случаев, причем в верхней строке записаны передаваемые сообщения, а в нижней — соответствующие принятые показания. Таблица позволяет определить любые необходимые вероятности, например одна верхняя строка дает априорное распределение вероятностей

$$p(\text{да}) = \frac{5}{8}, \quad p(\text{нет}) = \frac{3}{8}.$$

Однако, если появился «зеленый» сигнал, надо выбирать только из тех возможностей, которые соответствуют «зеленому» во второй строке. Среди них 3 из 4 соответствуют «да» и, следовательно, апостериорная вероятность для «да», когда принят «зеленый» сигнал, равна $\frac{3}{4}$. Это и есть применение «обратной вероятности».

Формально теорема устанавливается следующим образом. Пусть x — передаваемое сообщение, а y — принятое показание. Тогда по теореме об умножении вероятностей вероятность совмещения

$p(x, y)$, т. е. вероятность совместного появления значения x и значения y , равна

$$p(x, y) = p(x)p_x(y) = p(y)p_y(x), \quad (1)$$

где условная вероятность $p_x(y)$ есть вероятность появления y , если появился x .

Эти важные соотношения легко могут быть проверены при помощи табл. 1. Так, когда x есть «да», а y — «зеленый» сигнал, то

$$\frac{3}{8} = \frac{5}{8} \times \frac{3}{5} = \frac{4}{8} \times \frac{3}{4}.$$

Из пяти вероятностей, входящих в уравнение (1), две имеют особенное значение для теории передачи электрических сигналов, а именно $p(x)$ — априорная вероятность сообщения x и $p_y(x)$ — апостериорная вероятность сообщения x при условии приема сигнала y . При приеме в общем случае должны быть рассмотрены все значения x , так что $p(x)$ и $p_y(x)$ часто рассматриваются как распределения вероятностей. Их суммы по x равны, конечно, единице.

Теорема «обратной вероятности»¹⁾ представляет собой просто выражение для апостериорного распределения, полученное из (1), а именно:

$$p_y(x) = \frac{p(x)p_x(y)}{p(y)} \quad (2)$$

Предполагается, что наблюдатель знает априорные вероятности $p(x)$ и статистические свойства помех, определяемые $p_x(y)$, представляющим полное семейство распределений. При приеме сигнала y наблюдатель должен использовать уравнение (2) для определения относительных вероятностей того, что каждое из сообщений x было переданным сообщением. Так как y будет тогда фиксированным, $p(y)$ представляет постоянную, которая может быть вычислена, если $p_y(x)$ нормировано, т. е. его сумма по x равна единице.

В этом случае уравнение (2) может быть записано в виде

$$p_y(x) = kp(x)p_x(y), \quad (3)$$

где k — постоянная, не зависящая от x . Статистически $p_y(x)$ представляет относительную частоту, с которой x действительно передается в большом числе переданных сообщений, которые создают один и тот же сигнал y в приемнике.

Например, если принят «зеленый» сигнал:

$$p_{\text{зел.}}(\text{да}) = k \times \frac{5}{8} \times \frac{3}{5} = \left(\frac{3}{8}\right)k,$$

$$p_{\text{зел.}}(\text{нет}) = k \times \frac{3}{8} \times \frac{1}{3} = \left(\frac{1}{8}\right)k.$$

¹⁾ Эта теорема известна в теории вероятностей как теорема о вероятности гипотез, или теорема Бэйеса. (Прим. ред.)

Апостериорное распределение нормируется, если положить $k=2$

$$p_{\text{зел.}}(\text{да}) = \frac{3}{4}, \quad p_{\text{зел.}}(\text{нет}) = \frac{1}{4}$$

что уже было получено раньше. Так же точно, когда принят «красный» сигнал, апостериорное распределение будет

$$p_{\text{кр.}}(\text{да}) = \frac{1}{2}, \quad p_{\text{кр.}}(\text{нет}) = \frac{1}{2}.$$

Ясно, что в общем случае для каждого принятого сигнала будут различные апостериорные распределения вероятностей сообщений.

Во многих случаях различные принимаемые сигналы не дискретны и число их не конечно, а они образуют непрерывное множество. Это может быть, например, если y — напряжение в некоторый фиксированный момент времени. Тогда теорема «обратной вероятности» должна быть истолкована несколько иным образом.

Если y относится к непрерывному множеству, его распределение вероятностей будет характеризоваться непрерывной плотностью распределения, причем $p(y) dy$ — вероятность попадания y в интервал $(y, y+dy)$. При подстановке этих вероятностей в уравнение (2) дифференциалы dy сокращаются и уравнение (3) остается в прежней форме, где $p_x(y)$ будет плотностью вероятностей.

Формула также справедлива, если x непрерывно, так как тогда дифференциалы dx также сокращаются. Следовательно, теорема «обратной вероятности» применима в форме уравнения (3), когда x или y или, наконец, оба вместе являются непрерывными переменными, причем в этом случае используется соответствующая плотность вероятностей.

Как было показано, каждый возможный принятый сигнал приводит к своему собственному апостериорному распределению вероятностей для передаваемых сообщений. Если эти распределения усреднить в соответствии с вероятностями появления сигналов, а именно $p(y)$, то интересно отметить, что в результате получается априорное распределение вероятностей сообщений. Таким образом, интегрирование (или суммирование) уравнения (1) по y дает

$$\int p(y)p_y(x) dy = \int p(x)p_x(y) dy = p(x). \quad (4)$$

Слева здесь среднее апостериорного распределения, а справа — априорное распределение. В сокращенной записи среднее по всем значениям некоторой переменной z будет обозначаться Sr_z и уравнение (4) тогда запишется так:

$$Sr_y p_y(x) = p(x). \quad (5)$$

В нашем примере, как видно из табл. 1, два принимаемых сигнала — «зеленый» и «красный» — встречаются во всем множестве одинаково часто. Соответственно среднее с равными весами двух

апостериорных распределений ($3/4$, $1/4$) и ($1/2$, $1/2$) дает априорное распределение ($5/8$, $3/8$).

При применении «обратной вероятности» к задачам теории передачи сигналов первый основной шаг сводится к оценке $p_x(y)$, представляющем статистические свойства шумов в наиболее подходящем для решения задачи виде. Это часто оказывается значительно более сложным, чем может показаться из приведенных рассуждений, поскольку y не всегда является таким простым, как «зеленый» или «красный» сигнал. В общем случае y будет представлять полный сигнал, как, например, телеграфная посылка или отраженный сигнал радиолокационной станции, плюс шумы.

Однако при этом не возникают какие-либо принципиальные затруднения, поскольку можно рассматривать полный сигнал как последовательность ординат кривой напряжения и вычислить $p_x(y_1, y_2, y_3 \dots)$ как многомерное распределение вероятностей совмещения. Этот вопрос более детально рассмотрен в § 4.

2. ТЕОРИЯ ПОЛУЧЕНИЯ ДАННЫХ

Известно, что определенный выбор между двумя равновероятными событиями соответствует одной двоичной единице количества данных, а выбор между n равновероятными событиями — $\log_2 n$ двоичным единицам. В этом рассуждении предполагается полная определенность результата выбора, однако часто прием сигнала приводит только к изменению относительных вероятностей для ряда событий, без четкого выделения одного из них как достоверного. Поэтому необходимо разработать более общее определение, сводящееся к этому более простому, когда нет апостериорной неопределенности. Это и было сделано Шэнноном.

Настоящее изложение основано на работе Шэннона, но отличается своим подходом к задаче и исходными постулатами, поскольку вначале рассматривается действительно переданное количество данных, а не усредненное количество или средняя скорость передачи. Впрочем, определения Шэннона легко отсюда получаются, а некоторые его основные результаты кратко упоминаются в дальнейшем.

Начнем с двух следующих аксиом относительно аддитивности количества данных.

1. Если последовательно посылаются два сигнала, представляющие одно сообщение, а наблюдатель рассматривает апостериорную вероятность после приема первого сигнала как априорную вероятность перед приемом второго, то полное увеличение количества данных относительно этого сообщения равно сумме количества данных от каждого сигнала.

2. Если посылаются два сигнала, представляющие два независимых сообщения, то полное увеличение количества данных от-

носителем этих сообщений равно сумме количества данных, когда каждый сигнал рассматривается отдельно.

Из этих двух аксиом может быть развита вся дальнейшая математическая теория.

Обозначим сообщение, о котором идет речь в первой аксиоме, через x_i . Отметим три вероятности для этого сообщения: во-первых, его априорную вероятность $p(x_i)$, во-вторых, апостериорную вероятность $p_y(x_i)$ после приема первого сигнала и, в-третьих, принимая $p_y(x_i)$ за априорную вероятность для второго сигнала, обозначим конечную вероятность после приема второго сигнала через $p_z(x_i)$.

Допустим теперь, что после приема y и z наблюдатель вполне уверен, что было послано i -ое сообщение, так что $p_z(x_i) = 1$. Количество данных увеличилось, поскольку первоначальная неопределенность, связанная с $p(x_i)$, полностью устранена и полный выигрыш от приема y и z поэтому зависит только от $p(x_i)$. Аналогично увеличение количества данных от одного второго сигнала зависит только от $p_y(x_i)$. Соответственно увеличение количества данных от первого сигнала определяется только $p(x_i)$ и $p_y(x_i)$.

Таким образом, если остается неопределенность после того, как данное сообщение x_i было передано и принято при наличии шумов, количество данных является функцией априорной и апостериорной вероятностей одного этого сообщения. Оно может быть записано в виде

$$J[p(x_i), p_y(x_i)]. \quad (6)$$

Тогда первая аксиома утверждает

$$J[p(x_i), p_y(x_i)] + J[p_y(x_i), p_z(x_i)] \equiv J[p(x_i), p_z(x_i)]. \quad (7)$$

Как показано в Приложении, для удовлетворения этого тождества J должно иметь вид

$$J[p(x_i), p_y(x_i)] \equiv j[p(x_i)] - j[p_y(x_i)]. \quad (8)$$

Чтобы определить теперь функциональную форму j , необходимо воспользоваться второй аксиомой. Два независимых сообщения можно обозначить x_i и x_k , а соответствующие принятые сигналы y и z . Поскольку вероятность совмещения двух независимых событий равна произведению их вероятностей, вторая аксиома дает

$$\begin{aligned} j[p(x_i)p(x_k)] - j[p_y(x_i)p_z(x_k)] &\equiv \\ &\equiv j[p(x_i)] + j[p(x_k)] - j[p_y(x_i)] - j[p_z(x_k)]. \end{aligned} \quad (9)$$

Из этого тождества следует, как это показано в Приложении, что $j(p)$ должно иметь вид

$$j(p) = A \log p + B, \quad (10)$$

где постоянные A и B могут быть выбраны произвольно. Тогда из уравнений (6), (8) и (10) количество данных может быть записано в форме

$$J[p(x_i), p_y(x_i)] = -A \log \frac{p_y(x_i)}{p(x_i)}, \quad (11)$$

а чтобы увеличение вероятности истинного сообщения представляло положительное количество данных, A принимается равным -1 . Наконец, для краткости записи вместо J будем писать $I_{x,y}$, чтобы отметить, что это есть количество полученных данных, когда передано некоторое сообщение x , а принят некоторый сигнал y .

Тогда, пользуясь уравнением (1):

$$I_{x,y} = \log \frac{p_y(x)}{p(x)} = \log \frac{p(x,y)}{p(x)p(y)}. \quad (12)$$

Это основное выражение для количества данных в теории Шэннона.

К сожалению, когда принимается сигнал, искаженный шумами, наблюдатель не в состоянии полностью судить об истинном сообщении. Кроме того, с точки зрения наблюдателя, интуитивно чувствуется, что тот же самый принятый сигнал всегда представляет одно и то же количество данных независимо от действительно переданного сообщения.

Поэтому единственный способ определить увеличение количества данных для наблюдателя заключается в усреднении $I_{x,y}$ по всем случаям, в которых фиксирован один лишь сигнал y . Как показано в § 1, в этих условиях передаваемые сообщения появляются с относительными частотами, определяемыми $p_y(x)$. Тогда увеличение количества данных I_y для наблюдателя можно определить, усредняя $I_{x,y}$ и принимая $p_y(x)$ за весовой коэффициент:

$$I_y = \text{Cp}_x I_{x,y} = \sum_x p_y(x) \log \frac{p_y(x)}{p(x)}, \quad (13)$$

если сообщения дискретны.

Часто сообщения образуют континуум, как, например, в случае расстояний до самолета или показаний измерительного прибора. Распределение вероятностей x представляется тогда непрерывными кривыми плотности вероятностей, но теория без труда может быть распространена и на этот случай.

Если $p(x)$ — плотность распределения, то $p(x) \delta x$ есть вероятность того, что x лежит между x и $x + \delta x$. Если разбить область значений x на элементы шириной δx , то может быть использована дискретная теория и тогда, полагая $\delta x \rightarrow 0$:

$$I_y = \int p_y(x) \log \frac{p_y(x)}{p(x)} dx. \quad (14)$$

Следует отметить, что увеличение количества данных по уравнению (13) или (14) аддитивно в смысле аксиомы 2, но не в смысле аксиомы 1. Это объясняется тем, что увеличение количества данных для наблюдателя представляет апостериорное среднее, а усреднение после приема первого и второго сигналов производится по различным множествам.

Как и следовало ожидать, можно показать, что I_y (в противоположность $I_{x,y}$) никогда не является отрицательным и равно нулю только в том случае, когда распределения $p(x)$ и $p_y(x)$ одинаковы, так как тогда прием сигнала совершенно не изменяет состояния знания наблюдателя.

Полученные выше соотношения могут быть иллюстрированы применением их к примеру, приведенному в § 1. Если послано сообщение «да» и принят «зеленый» сигнал, то переданное количество данных определяется из уравнения (12):

$$\log p_{\text{зел.}}(\text{да}) - \log p(\text{да}) = \log \frac{3}{4} - \log \frac{5}{8} = \log \frac{6}{5}.$$

Аналогично, если послано «нет» и опять принят «зеленый» сигнал, значение $I_{x,y}$ равно

$$\log p_{\text{зел.}}(\text{нет}) - \log p(\text{нет}) = \log \frac{1}{4} - \log \frac{3}{8} = \log \frac{2}{3}$$

Оно отрицательно, потому что вероятность сообщения «нет», которое в действительности было послано, уменьшилась в результате приема сигнала.

Ни одно из приведенных выше выражений не представляет, однако, большой ценности для наблюдателя, поскольку единственное его знание относительно того, что было послано, дается апостериорным распределением

$$p_{\text{зел.}}(\text{да}) = \frac{3}{4} \quad p_{\text{зел.}}(\text{нет}) = \frac{1}{4}.$$

Поэтому увеличение количества данных для наблюдателя равно

$$I_{\text{зел.}} = \frac{3}{4} \log \frac{6}{5} + \frac{1}{4} \log \frac{2}{3} = 0,0510 \text{ двоичной единицы,}$$

что является примером применения уравнения (13). Так же точно при приеме «красного» сигнала

$$I_{\text{кр.}} = \frac{1}{2} \log \frac{4}{5} + \frac{1}{2} \log \frac{4}{3} = 0,0466 \text{ двоичной единицы.}$$

Как $I_{\text{зел.}}$, так и $I_{\text{кр.}}$ обязательно положительны.

Вычисленные выше количества данных относятся к определенным принятым сигналам, но когда необходимо найти пропускную способность канала, важное значение имеет среднее количество данных I , когда ни x , ни y не заданы. Эта величина применялась Шэнноном. Она получается путем усреднения количества данных,

полученных при приеме каждого сигнала, причем этим данным придан вес в соответствии с вероятностью появления сигналов.

В нашем примере «красный» и «зеленый» сигналы встречаются одинаково часто (см. табл. 1) и поэтому

$$I = \frac{1}{2} (0,0510 + 0,0466) = 0,0488 \text{ двоичной единицы.}$$

Общее выражение получается, если применить оператор C_{p_y} к уравнению (14):

$$I = C_{p_y} I_y = \int p(y) \int p_y(x) \log \frac{p_y(x)}{p(x)} dx dy. \quad (15)$$

Это то же самое, что и среднее от $I_{x,y}$ по всем x и по всем y , которое по уравнению (12) имеет симметричную форму:

$$I = C_{p_{x,y}} I_{x,y} = \iint p(x,y) \log \frac{p(x,y)}{p(x)p(y)} dx dy. \quad (16)$$

Еще другая форма I получается в том случае, если сначала разделить логарифм в уравнении (15) на две части

$$I = C_{p_y} \int p_y(x) \log p_y(x) dx - \iint p(y) p_y(x) \log p(x) dx dy.$$

Интегрирование по y может быть выполнено при помощи уравнения (4); тогда

$$I = H(x) - H_y(x), \quad (17)$$

где

$$H(x) = - \int p(x) \log p(x) dx, \quad (18)$$

$$H_y(x) = - C_{p_y} \int p_y(x) \log p_y(x) dx. \quad (19)$$

Шэннон называет $H(x)$ «энтропией» распределения $p(x)$, и уравнение (17) тогда утверждает, что среднее количество данных, приходящихся на сигнал, равно разности априорной и апостериорной «энтропий». Из симметрии уравнения (16) следует, что I может быть также написано в форме

$$I = H(y) - H_x(y) \quad (20)$$

путем перемены местами x и y .

Из этого последнего выражения Шэннон получает свою теорему 17 — один из основных результатов теории передачи сигналов. Она утверждает, что среднее количество данных, которое может быть передано за время T по каналу с полосой частот W при наличии гауссовых шумов с равномерным спектром средней мощностью N , может достигать, но никогда не превосходит

$$I_{\text{макс.}} = WT \log \left(1 + \frac{P}{N} \right), \quad (21)$$

где P — средняя мощность принимаемых сигналов.

Эта теорема показывает ограниченность пропускной способности канала. Обычно мощность шумов возрастает пропорционально полосе частот: $N = WN_0$, где N_0 — средняя мощность шумов в единице полосы частот. Тогда $I_{\text{макс.}}$ растет вместе с W , стремясь в пределе к

$$\lim_{W \rightarrow \infty} I_{\text{макс.}} = \frac{PT}{N_0} = \frac{E}{N_0} \text{ натуральных единиц,} \quad (22)$$

где E — полная энергия сигналов.

Уравнения (21) и (22) в принципе позволяют измерить эффективность любой системы передачи сигналов как при наличии, так при отсутствии ограничения величины полосы частот канала. Конечно, чрезвычайно затруднительно оценить таким способом эффективность всей системы, предназначенной, например, для передачи музыки. Однако такие системы, как радиолокационные станции, линии радиосвязи с кодовой импульсной модуляцией и с временной импульсной модуляцией, сравнительно просто поддаются анализу.

Можно показать, что при кодовой импульсной модуляции и временной импульсной модуляции необходима мощность сигнала примерно на 8 дБ больше, чем в идеальной системе, удовлетворяющей уравнению (21). Вместе с тем доказано¹⁾, что в лучшем случае при измерении расстояний радиолокационной станцией достигается весьма близкое приближение к идеальной системе, определяемой уравнением (22).

Все полученные соотношения одинаково применимы как к случаю дискретных распределений вероятностей, так и к случаю непрерывных распределений, если только заменить суммы интегралами. Это связано с тем существенным обстоятельством, что под знак логарифма входят только отношения вероятностей (либо плотностей вероятностей).

Путем простого постулирования аддитивности количества данных была показана возможность измерения количества данных как логарифмического изменения вероятностей. Формальное определение $I_{x,y}$ в уравнении (12) непосредственно имеет мало значения, поскольку оно предполагает знание как переданного сообщения x , так и принятого сигнала y . Когда наблюдатель принимает сигнал, его более интересует значение $I_{x,y}$, усредненное по всем сообщениям, которые могли быть переданы, т. е. I_y . Наконец, I_y может быть усреднено по всем сигналам y , чтобы получить I — среднее количество данных, приходящихся на сигнал.

3. ШУМЫ И АПОСТЕРИОРНОЕ РАСПРЕДЕЛЕНИЕ

Чтобы применить излагаемую теорию к практическим задачам, необходимо рассмотреть, как может быть построено апостериорное

¹⁾ См. следующую главу. (Прим. ред.)

распределение вероятностей по принятому сигналу y . Предполагается, что априорное распределение $p(x)$ известно и тогда апостериорное распределение $p_y(x)$ может быть сразу найдено из уравнения (3), как только вычислено $p_x(y)$.

Условная вероятность $p_x(y)$ описывает действие шумов на систему. Она определяет случайный характер принимаемого сигнала при фиксированном передаваемом сообщении. Если принимаемый сигнал $y(t)$ состоит из исходного сигнала $u_x(t)$, представляющего сообщение x , и гауссовых шумов с равномерным спектром, вычисление $p_x(y)$ может быть непосредственно выполнено, как это показано ниже.

Если $u_x(t)$ — сигнал, который был бы в отсутствие шумов, распределение вероятностей для принятого сигнала $y(t)$ (при действии помех) имеет вид

$$p_x(y) = G(y - u_x), \quad (23)$$

где $G(n)$ — плотность вероятностей для шумов $n(t)$.

Простой способ точного вычисления G основан на известной теореме¹⁾ о том, что если функция времени $f(t)$ не содержит частот, превышающих W , то

$$f(t) \equiv \sum_r f\left(\frac{r}{2W}\right) \frac{\sin \pi(2Wt-r)}{\pi(2Wt-r)}. \quad (24)$$

Важным свойством функции $\frac{\sin \pi x}{\pi x}$ является то, что она равна нулю, когда x — целое число, и единице, когда $x = 0$, а также, что

$$\int_{-\infty}^{\infty} \frac{\sin \pi x}{\pi x} dx = 1$$

и

$$\int_{-\infty}^{\infty} \frac{\sin \pi(x-r)}{\pi(x-r)} \cdot \frac{\sin \pi(x-s)}{\pi(x-s)} dx = \begin{cases} 1, & r = s \\ 0, & r \neq s, \end{cases}$$

где r и s — целые числа.

Чтобы применить эту теорему, необходимо предположить, что все рассматриваемые колебания пропущены через фильтр нижних частот, коэффициент передачи которого постоянен вплоть до частоты W и равен нулю на всех более высоких частотах. Если выбрать граничную частоту W достаточно большой, то сигнал останется неизменным и никакой потери общности это предположение не вызовет. Позже будет показано, что точная величина W не имеет значения и W исчезает из последующих формул.

¹⁾ См. часть I, теорема 13. (Прим. ред.)

Рассмотрим сначала гауссовы шумы с равномерным спектром $n(t)$. На выходе фильтра каждое значение $n\left(\frac{r}{2W}\right)$ в точке отображения¹⁾, обозначаемое в дальнейшем n_r , имеет по определению распределение вероятностей

$$p(n_r) = \exp\left(-\frac{n_r^2}{2N}\right), \quad (25)$$

где N — средний квадрат $n(t)$ или средняя мощность шумов.

Можно показать, что эти значения шумов в точках отображения статистически независимы, если только спектр шумов равномерен в пределах всей полосы частот вплоть до W гц. Следовательно, распределение вероятностей совмещения для всего множества «дискрет» шумов равно произведению отдельных распределений. Поскольку «дискрет» определяют рассматриваемые колебания шумов, то это произведение дает плотность вероятностей для самих колебаний шумов. Поэтому

$$G(n) = \exp\left(-\sum \frac{n_r^2}{2N}\right) \quad (26)$$

Возводя в квадрат основное тождество (24), интегрируя по времени и пользуясь свойствами функции $\frac{\sin \pi x}{\pi x}$, сумму под знаком показательной функции можно представить интегралом, и тогда

$$G(n) = \exp\left[-\frac{1}{N_0} \int n^2(t) dt\right]. \quad (27)$$

Здесь N_0 есть средняя мощность шумов в единице полосы частот, имеющая размерность энергии и являющаяся основным параметром шумов.

Теперь апостериорное распределение для сообщения x может быть написано в явной форме. Из уравнений (3), (23) и (27)

$$p_y(x) = k p(x) \exp\left[-\frac{1}{N_0} \int (y - u_x)^2 dt\right]. \quad (28)$$

Это соотношение является основным вероятностным уравнением во всех задачах, когда помехи создаются только за счет добавления к сигналу гауссовых шумов с равномерным спектром.

Постоянная k выбирается из условий нормировки $p_y(x)$; интеграл под знаком показательной функции берется по всему промежутку времени, в течение которого производится передача сигнала.

¹⁾ Т. е. значения колебаний шумов, рассматриваемые в дискретных точках, отстоящих друг от друга по оси времени на интервал $t_0 = \frac{1}{2W}$. В дальнейшем применяется термин «дискрет» шумов. (Прим. ред.) $\frac{1}{2W}$

Отсюда следует, что если отвлекаться от априорного весового коэффициента, то наиболее вероятным сообщением является такое сообщение, форма колебаний которого $u_x(t)$ имеет наименьшее эффективное отклонение от формы принимаемых колебаний $y(t)$.

4. КОРРЕЛЯЦИОННЫЙ ПРИЕМНИК

Может показаться, что описанные математические операции полезны только при вычислении количества данных, несомых сигналом, однако вероятностное уравнение (28) имеет более широкое значение.

В принципе оно указывает, как должен быть построен идеальный приемник для извлечения всего наличного количества данных из смеси сигнала и шумов на входе приемника.

Идея, на которой основывается это утверждение, теперь должна быть очевидной, поскольку $p_y(x)$ само является наличным количеством данных. Разумные требования, которые можно предъявить к приемнику, должны исходить из условия, чтобы на основании выходных сигналов наблюдатель мог оценить относительные вероятности того, что каждое возможное сообщение является истинным. Если приемник позволяет вычислить эти вероятности, то никакого другого истолкования принимаемых сигналов не требуется. Рассматриваемое с этой точки зрения уравнение (28) математически описывает свойства идеального приемника.

Теперь будут обсуждены следствия этого утверждения. Допустим, что интеграл под знаком показательной функции записан в виде

$$\int y^2(t) dt - 2 \int y(t) u_x(t) dt + \int u_x^2(t) dt.$$

При приеме y^2 фиксировано и поэтому не зависит от x , которое можно рассматривать как испытываемое сообщение. Поэтому первый интеграл является постоянным множителем в $p_y(x)$ и может быть включен в k .

Во многих задачах интеграл от u_x^2 также не зависит от x , вследствие того что сигналы, представляющие различные сообщения, все имеют одинаковую энергию. Если это так, третий член также можно исключить, и уравнение (28) тогда превращается в

$$p_y(x) = kp(x) \exp \left[\frac{2q(x)}{N_0} \right], \quad (29)$$

где

$$q(x) = \int y(t) u_x(t) dt. \quad (30)$$

Начинать вычисления нужно с образования q путем поочередного умножения функции $y(t)$, описывающей принимаемый сигнал,

на функции $u_x(t)$, представляющие все возможные исходные сигналы, и последующего интегрирования по времени. $q(x)$ является основным членом в уравнении (29) и представляет собой оценку корреляции между y и u_x . Функция $q(x)$ возрастает, если $u_x(t)$ —сигнал, соответствующий действительно переданному сообщению.

Операция образования q из y обычно является необратимой. При таких процессах в общем случае происходит потеря данных, однако рассматриваемая операция связана с потерей только ненужных данных относительно шумовых составляющих сигнала $y(t)$.

Если в конечном итоге требуется определить наиболее вероятное сообщение, а априорные вероятности сообщений все равны, достаточно вычислить только $q(x)$. Остальная часть уравнения (29) представляет просто амплитудные искажения $q(x)$. Поскольку искажения монотонны, то в результате значение x , при котором наступает максимум $p_y(x)$, обеспечивает также максимум $q(x)$.

Рассмотрим теперь влияние множителя $\frac{2}{N_0}$ под знаком показательной функции. Этот множитель действует таким образом, что, когда шумы малы, показательная функция значительно усиливает изменения $q(x)$ с x . Именно это и следовало ожидать, так как если шумы невелики, то должно быть мало сомнений относительно истинного сообщения. Поэтому кривая $p_y(x)$ должна иметь ясно выраженный максимум, представляющий высокую степень надежности.

В другом крайнем случае, когда шумы настолько велики, что полностью перекрывают сигнал, показательная функция за счет множителя $\frac{2}{N_0}$ становится весьма малой по сравнению с единицей.

Поэтому $p_y(x)$ просто совпадает с $p(x)$ и уравнение (14) показывает, что при приеме сигнала количество данных не увеличивается.

Как будет видно, промежуточное положение возникает не тогда, когда средняя мощность сигнала P примерно равна мощности шумов N , а в том случае, когда полная энергия сигнала E сравнима с мощностью шумов в единице полосы частот N_0 . Действительно, P может быть много меньше N и в этом заключается большое преимущество корреляционного приемника. Это также видно непосредственно из уравнения (22), где именно отношение $\frac{E}{N_0}$ определяет максимальное количество данных в идеальной системе. Впрочем, не следует думать, что корреляционный метод обеспечивает нечто большее, чем дает интегрирование при приеме сигналов.

На практике, повидимому, корреляционные методы могут быть использованы только в простейших системах. В наших рассуждениях всюду молчаливо подразумевалось, что сигналы $u_x(t)$, соответствующие различным сообщениям, точно известны перед приемом.

В радиосвязи дело осложняется тем, что начальное временное положение сигнала заранее неизвестно.

В этом случае теория усложняется, так как функция сигнала, соответствующего сообщению x , должна быть записана в виде $u_x(t-\tau)$, где τ — неизвестное временное запаздывание. Апостериорное распределение тогда становится интегралом, взятым по всем возможным временным запаздываниям:

$$p_y(x) = k \int p(x, \tau) \exp \left[\frac{2}{N_0} \int y(t) u_x(t - \tau) dt \right] d\tau. \quad (31)$$

Это означает, что прежде всего должна быть установлена взаимная корреляция во все возможные моменты времени между принятыми сигналами и всеми возможными сигналами, представляющими сообщения. Во всех системах, за исключением простейших, это приведет к недопустимым на практике усложнениям.

В случае радиолокации сообщение (дальность до цели) представляется самым временным запаздыванием, а другие параметры колебаний постоянны. Для простоты здесь предполагается, что имеется только одна цель и сила отраженных сигналов известна и не зависит от дальности.

Тогда апостериорное распределение для временного запаздывания будет

$$p_y(\tau) = kp(\tau) \exp \left[\frac{2}{N_0} \int y(t) u(t - \tau) dt \right]. \quad (32)$$

В этой формуле $y(t)$ и $u(t)$ — высокочастотные колебания и, оставляя в стороне вопрос о пределах интегрирования, рассмотренный в другой работе ¹⁾, интеграл имеет форму выходного эффекта линейного фильтра. На входе действуют принимаемые колебания $y(t)$, а импульсная реакция есть $u(-t)$, т. е. функция, обратная по времени передаваемым колебаниям.

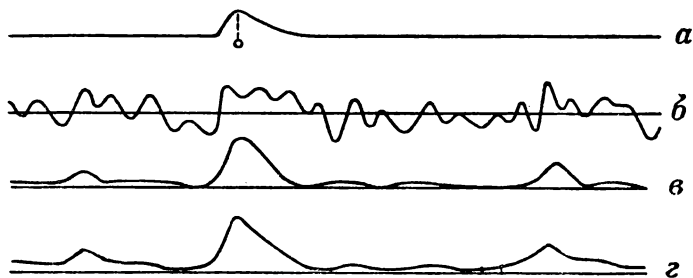
Подобный фильтр устанавливает взаимную корреляцию y и u по высокой частоте, т. е. выполняет ту же самую операцию, что и обычный приемник. Необходимость детектирования из уравнения (32) не очевидна, поскольку теоретически в этом нет необходимости при определении расстояния до неподвижной цели. Детектирование только уничтожает тонкую структуру сведений о дальности до цели, получаемую от несущей. Дальнейшее рассмотрение уравнения (32) более полно проделано в другой работе ¹⁾.

Для иллюстрации уравнения (32) в случае, когда все колебания низкочастотные, был выполнен простой эксперимент. На фиг. 1,а изображен передаваемый сигнал, а числовое сообщение x предполагается закодированным таким образом, что время запаздывания τ этого сигнала равно x . На фиг. 1,б показаны типичные колебания y , появляющиеся при добавлении шумов, отфильтро-

¹⁾ См. следующую главу. (Прим. ред.)

ванных в некоторой произвольной полосе W , как об этом говорилось раньше.

После того как наблюдателю было указано, что импульс может находиться с равной вероятностью в любом месте линии развертки (априорная оценка), ему было предложено на основании одного лишь наблюдения установить апостериорное распределение вероятностей для временного положения импульса. Наблюдатель был ознакомлен с формой и амплитудой импульса, но ему было неизвестно истинное положение отраженного сигнала.



Фиг. 1. Экспериментальное определение положения импульса.

a — сигнал; $б$ — сигнал и шум; $в$ — апостериорное распределение по субъективной оценке наблюдателя; $г$ — вычисленное апостериорное распределение.

Его субъективная кривая изображена на фиг. 1, $в$, а теоретическая кривая, вычисленная по уравнению (32), показана на фиг. 1, $г$. Если отношение сигнал/шум возрастает, то кривая фиг. 1, $г$ стремится к дельта-функции.

Идеальный приемник должен осуществлять преобразование кривой фиг. 1, $б$ в кривую фиг. 1, $г$ чисто электронными устройствами, выдавая относительные вероятности всех возможных сообщений, благодаря чему отпадает необходимость в субъективной оценке.

5. ЗАКЛЮЧЕНИЕ И ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Было показано, что из всех распределений вероятностей, которые могут быть использованы для описания статистических свойств системы передачи сигналов, два распределения имеют особое значение. Это — распределения вероятностей различных возможных «сообщений» перед и после приема сигнала. Процесс приема может рассматриваться как событие, изменяющее относительные вероятности сообщений, а теория получения данных обеспечивает числовую меру этого изменения.

Эти идеи естественно приводят к представлению о том, что весь процесс приема есть просто способ определения или выявления апостериорных вероятностей сообщений. Следовательно,

выражение для апостериорного распределения само представляет математическое описание свойств оптимального приемника.

Как можно заметить, во всех уравнениях для апостериорной вероятности априорное распределение появляется как весовой коэффициент. Это основное свойство «обратной вероятности» может привести к затруднениям. В статистике часто оказывается, когда делается попытка произвести выбор между различными гипотезами (сообщениями) в свете некоторых новых данных (сигнала), что нет очевидного способа априорной оценки гипотез.

Хотя это затруднение вряд ли может возникнуть в системе связи, оно часто появляется в системах наблюдения, таких, как радиолокация. В этом случае при обстоятельствах, совершенно отличных от тех, с которыми мы до сих пор встречались, может оказаться необходимым приписать дальности до цели некоторое априорное распределение вероятностей. Если нет статистических сведений, на которых можно было бы основывать априорное распределение, то как оно вообще может быть определено?

Здесь иногда представляется заманчивым постулировать некоторую неопределенную функцию просто как «формальный способ выражения незнания». В лучшем случае это — несколько произвольная процедура, и возражения против нее хорошо выражены Бартлеттом¹⁾, который пишет: «Подстановка простой функции для априорной вероятности, которая, если она вообще могла бы быть вычислена, определенно потребовала бы формулировки всех исходных данных, дает апостериорную вероятность в точной форме, которая, однако, может привести к серьезным заблуждениям. Кроме того, в этой попытке сделать конечный вывод относительно параметра (сообщения) и дать точную вероятность каждого возможного значения, мы вынуждены смешать сведения, которые могут быть извлечены из образца (сигнала), с другими сведениями, возможно, имеющимися в нашем распоряжении».

Необходимо различать два пути, по которым происходит использование принципа «обратной вероятности» в этой работе. Во-первых, он используется в процессе измерения увеличения количества данных в смысле Шэннона. Хотя этот прирост данных определяется самим сигналом, его величина оказалась зависящей от априорного распределения. Если нельзя построить априорного распределения, то невозможно определить, в какой степени сигнал просто дублирует существующее знание. Это, конечно, не оправдывает использования чисто субъективных априорных распределений в теории Шэннона, которая основана на частотном определении вероятности, но оправдывает применение априорных множителей всякий раз, когда имеются предварительные статистические данные.

¹⁾ М. С. Бартлетт, «Вероятность и случайность в теории статистики», *Proc. of the Royal Soc. A*, 141, 518, 1933.

Во-вторых, в работе предлагается использовать «обратную вероятность» для описания свойств идеального приемника. Можно разъяснить это положение. Нет необходимости требовать, чтобы идеальный приемник выдавал вычисленное до конца апостериорное распределение, поскольку последнее легко находится по выходному сигналу. Таким образом, вызывающий затруднения априорный множитель вполне может быть исключен из математического описания свойств идеального приемника, когда этот множитель вызывает сомнение и на практике вводится субъективной оценкой наблюдателя.

Может возникнуть вопрос, почему необходимо требовать, чтобы в идеальном приемнике производилось определение апостериорных вероятностей каждого возможного сообщения, а не просто выделялось наиболее вероятное сообщение, поскольку на практике обычно только это и нужно. Действительно, многие системы связи чрезвычайно бы усложнились, если бы на выходе приемника требовалось определять вид апостериорного распределения (возможно, многомерного). Однако нет необходимости всегда так буквально понимать развиваемую теорию.

Апостериорное распределение может по желанию рассматриваться как средство определения наиболее вероятного сообщения. Тогда может быть построен приемник, выдающий на выходе те значения x , при которых $p_y(x)$ достигает максимума.

Иногда имеются возражения против такой процедуры. Во-первых, сила критики Бартлетта значительно возрастает, поскольку априорное распределение используется необратимым способом. Во-вторых, достоверность сообщения перестает быть очевидной наблюдателю. И, наконец, в некоторых случаях, как, например, в радиолокации, после некоторого наблюдения могут поступать еще сигналы, относящиеся к тому же сообщению. При этом апостериорное распределение после первоначального наблюдения необходимо как априорное распределение для следующего сигнала.

Как показано в другой работе ¹⁾, именно так обстоит дело, когда в радиолокационной системе осуществляется суммирование сигналов от импульса к импульсу. Преждевременный выбор оптимального сообщения делает невозможным такое комбинирование нескольких сигналов, и в результате может произойти потеря части полезных сведений.

Приложение

Функция $J[\rho(x_i), p_y(x_i)]$ может быть написана в более простой форме $J(\xi, \eta)$ как функция двух переменных ξ и η . Тогда тождество $\underline{1}(7)$ утверждает, что

$$J(\alpha, \beta) + J(\beta, \gamma) \equiv J(\alpha, \gamma), \quad (33)$$

¹⁾ См. работу Вудворда в главе III. (Прим. ред.)

где α , β и γ — частные значения переменных ξ и η . Предполагается, что $J(\xi, \eta)$ дифференцируема по ξ . Тогда, рассматривая в уравнении (33) приращение α до $\alpha + \delta\alpha$:

$$\left[\frac{\partial J(\xi, \eta)}{\partial \xi} \right]_{(\alpha, \beta)} \equiv \left[\frac{\partial J(\xi, \eta)}{\partial \xi} \right]_{(\alpha, \gamma)}$$

Поскольку это тождество справедливо для всех значений β и γ , то $\frac{\partial J(\xi, \eta)}{\partial \xi}$ не зависит от η . Поэтому интегрирование по ξ дает

$$J(\xi, \eta) \equiv j(\xi) + k(\eta), \tag{34}$$

где $j(\xi)$ не зависит от η , а $k(\eta)$ не зависит от ξ . Но при подстановке из уравнения (34) в уравнение (33)

$$j(\cdot) \equiv -k(\beta)$$

и, следовательно:

$$J(\xi, \eta) \equiv j(\xi) - j(\eta). \tag{35}$$

Уравнение (9) может быть теперь записано в форме

$$j(\alpha') - j(\gamma\delta) \equiv j(\alpha) + j(\xi) - j(\gamma) - j(\delta).$$

Поддерживая γ и δ постоянными, напомним

$$j(\alpha') \equiv j(\alpha) + j(\cdot) + \text{const.} \tag{36}$$

Если предположить, что $j(\xi)$ дифференцируема, то, рассматривая небольшие изменения α в уравнении (36), найдем

$$\beta \left[\frac{dj(\xi)}{d\xi} \right]_{(\alpha\beta)} \equiv \left[\frac{dj(\xi)}{d\xi} \right]_{(\alpha)}$$

Полагая $\alpha = 1$:

$$\left[\frac{dj(\xi)}{d\xi} \right]_{(\beta)} \equiv \frac{A}{\beta}, \tag{37}$$

где A — постоянная, равная $j'(1)$. Поскольку уравнение (37) — тождество, то

$$j'(\xi) = \frac{A}{\xi}$$

и, следовательно:

$$j(\xi) = A \lg \xi + B, \tag{38}$$

где A и B — произвольные постоянные. Это и есть искомым результатом; использованный в уравнении (10).

Г л а в а II

СТАТИСТИЧЕСКАЯ ТЕОРИЯ ПРИЕМА РАДИОЛОКАЦИОННЫХ СИГНАЛОВ ¹⁾

Ф. ВУДВОРД и И. ДЭВИС

1. ВВЕДЕНИЕ

Известно, что количество данных, которое может быть получено при любом физическом наблюдении, ограничено флуктуационными процессами. По этой причине в радиотехнике всегда стремятся работать с сигналами, большими сравнительно с внутренними шумами канала. На первый взгляд можно заключить, что сообщения могут передаваться с удовлетворительной скоростью только в том случае, если мощность сигнала P больше мощности внутренних шумов в системе N . Однако столь резкой границы не существует, ибо, как показано в части I, предельная скорость, с которой могут передаваться сообщения в присутствии шумов с равномерным спектром, равна

$$C = W \log_2 \left(1 + \frac{P}{N} \right), \quad (1)$$

если полоса ограничена частотой W . Тем не менее во многих практических системах при $P=N$ обнаруживается порог разборчивости, однако это связано только со способом кодирования сообщений. Это не следует из уравнения (1), так как здесь подразумевается идеальное кодирование, которое может усложниться при $P < N$. Порог особенно заметен, если при кодировании происходит нетопологическое возрастание первоначального числа измерений ²⁾.

¹⁾ Ph. M. Woodward and I. L. Davies, «A Theory of Radar Information», *The Philosophical Magazine*, 41, № 321, 1001—1017, October 1950.

²⁾ Как показано В. А. Котельниковым (см. § 18, часть I), множество функций времени со спектром, ограниченным полосой частот W , и определенных на интервале T может быть представлено совокупностью точек в пространстве $2WT$ измерений. При такой геометрической трактовке процесс кодирования эквивалентен преобразованию «пространства сообщений» в «пространство сигналов». При этом преобразовании размерность пространства сигналов, вообще говоря, может стать отличной от размерности пространства сообщений. При топологическом преобразовании, т. е. когда преобразование осуществляется однозначным непрерывным образом, размерности пространства должны быть одинаковы. При нетопологическом преобразовании, например, в случае передачи с частотной модуляцией или с кодовой импульсной модуляцией, размерность пространства сигналов должна быть существенно большей. В таких системах пороговый эффект действия помех выражен весьма заметно. (Прим. ред.)

Простейшие данные, выдаваемые радиолокационной системой, — расстояния до отражающего объекта, который в этой работе считается неподвижным. Такие данные обладают только одним измерением, но так как они кодируются специальным образом, то размерность возрастает. В процессе работы повторяющиеся колебания (не обязательно последовательность импульсов) излучаются, отражаются объектом и принимаются. Предполагается, что наблюдение продолжается только конечное время — некоторое целое число периодов повторения, благодаря чему приемник получает конечное значение энергии сигнала. Наблюдатель сравнивает передаваемые и принимаемые колебания и старается определить время запаздывания. Наличие шумов при приеме ограничивает точность наблюдения.

Как передаваемые, так и принимаемые колебания занимают значительную площадь на частотно-временной диаграмме и поэтому заключают в себе значительное число элементарных ячеек ¹⁾. Пользуясь геометрической трактовкой ²⁾, можно сказать, что их размерность высока. Но так как форма передаваемых колебаний неизменна и вовсе не содержит нужных сведений, а принимаемые колебания представляют собой запаздывающую копию передаваемых, то неясно, можно ли также успешно использовать эту площадь на частотно-временной диаграмме, как и в обычной системе связи. Одна из задач работы — показать, что в этом смысле радиолокационные системы могут быть весьма эффективными.

Основная задача заключается, однако, в том, чтобы определить порог разборчивости ³⁾, который можно ожидать резко выраженным из-за принятого метода кодирования. В условиях разборчивого приема возможно определить точность, с которой может быть установлено значение расстояния до цели.

Математический анализ сосредоточивается вокруг одного важного понятия — распределения апостериорных вероятностей для времени запаздывания τ между передаваемым и принимаемым сигналами. Это распределение вероятностей описывает все, что может быть известно относительно расстояния, после того как было

¹⁾ Понятие об элементарной ячейке $\Delta f \Delta t$ (или об элементарном сигнале) связано со следующим. Для того чтобы сигнал, ограниченный полосой частот Δf , можно было зарегистрировать на выходе приемного устройства, он должен воздействовать на это устройство некоторый минимальный промежуток времени Δt . В результате можно говорить об элементарной ячейке площадью $\Delta f \Delta t$ на частотно-временной диаграмме. Подробнее см. А. А. Харкевич, «Спектры и анализ», Гостехтеоретиздат, глава 1, §13, 1952. (Прим. ред.)

²⁾ См. примечание на стр. 258. (Прим. ред.)

³⁾ В дальнейшем мы сохраняем термины «порог разборчивости» и «разборчивый прием», заимствованные из техники радиосвязи, хотя следует помнить, что в данном случае речь идет о зрительном наблюдении, а не о слуховом приеме речи. (Прим. ред.)

произведено наблюдение. Оно формально выводится в § 3 после предварительного введения необходимой терминологии в § 2.

Как оказывается, распределение зависит от двух функций τ , для которых приняты термины «функция сигнала» и «функция шумов». Они представляют собой функцию автокорреляции сигнала и функцию взаимной корреляции сигнала и шумов. Они обе зависят от известной формы передаваемых колебаний и от известных статистических характеристик шумов, но, кроме того, в отдельности зависят и от действительно принимаемых сигналов и шумов.

Сравнительное значение этих функций обсуждается в § 4, где доказываются некоторые их свойства. Так называемая функция сигнала показывает, в какой степени ограничена точность определения расстояния в присутствии шумов (§ 5). Ограничение точности, которое она описывает, представляет собой просто «топологическую неопределенность» наблюдения, поскольку она относится к области неопределенности в окрестности истинного значения дальности.

Вместе с тем функция шумов может создать ложные пики в распределении вероятностей, представляющие «нетопологическую неопределенность» или ненадежность, которая может появиться в результате наблюдения и сделать наблюдение бесполезным с практической точки зрения. Эта ненадежность возникает из-за нетопологического преобразования, упоминавшегося раньше.

Условия, при которых в радиолокационной системе возникает ненадежность, рассматриваются в § 6, где выводится формула для порога разборчивости. Наконец, в § 7 находится «энтропия» апостериорного распределения. За вычетом априорной «энтропии» она описывает количество данных, полученных при наблюдении. Интересно сравнить этот результат с общей формулой для пропускной способности канала связи. Полученные результаты кратко обсуждаются в заключительном параграфе.

2. МАТЕМАТИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ КОЛЕБАНИЙ

Передаваемые высокочастотные колебания будут описываться комплексной функцией $\psi(t)$, а принимаемые колебания — функцией

$$\gamma(t) = \alpha\psi(t - \tau) + \nu(t). \quad (2)$$

При приеме $\psi(t)$ претерпевает изменения по амплитуде и запаздывает на интервал τ при прохождении до цели и обратно. $\nu(t)$ есть комплексная случайная функция, представляющая шум. Все эти комплексные функции образованы из вещественных функций, описывающих колебания, путем добавления мнимой составляющей в квадратуре по методу, изложенному, например, в работе Гэбора¹⁾.

¹⁾Д. Гэбор, *Journ. Inst. Electr. Eng.*, 93, ч. III, 429; 1945.

Проведение анализа при помощи таких комплексных функций не приводит к каким-либо ошибочным выводам; необходимо только помнить, что энергия, рассеиваемая в единичном сопротивлении, теперь не равна интегралу от квадрата функции, а составляет половину интеграла от квадрата модуля функции.

Ради чисто математических удобств будем полагать, что энергия шумов сосредоточена в конечной полосе частот. Если эта полоса достаточно широка, чтобы включить весь спектр $\psi(t)$, то никакой потери общности такое предположение не вносит. Действительно, оно просто эквивалентно допущению, что принимаемые колебания $\gamma(t)$ профильтрованы таким образом, чтобы ослабить шумы, оставив сигнал неизменным. Подобное ограничение шумов, предполагаемых равномерно распределенными по частоте, не имеет отношения к задаче об оптимальной фильтрации.

Поскольку γ , ψ и ν являются высокочастотными функциями, их можно написать в виде

$$\begin{aligned}\gamma &= y(t) e^{2\pi i f_0 t}, \\ \psi &= u(t) e^{2\pi i f_0 t}, \\ \nu &= n(t) e^{2\pi i f_0 t},\end{aligned}\tag{3}$$

где y , u и n — комплексные низкочастотные функции, а f_0 — несущая частота, определяемая из

$$f_0 = \frac{1}{2\pi i} \frac{\int \psi^* \frac{d}{dt} \psi dt}{\int \psi^* \psi dt}.\tag{4}$$

Можно показать, что при таком определении f_0 является центральной частотой энергетического спектра сигнала ψ . Когда функция ψ периодична, то это эквивалентно условию

$$\int |\dot{\psi}|^2 (f - f_0) dt = 0,\tag{5}$$

где интеграл берется по одному периоду, а f обозначает мгновенную частоту или скорость изменения фазы колебаний $\psi(t)$. Приведенные определения обладают тем преимуществом, что они одинаково хорошо применимы к любому виду модуляции.

В дальнейшем удобно ввести векторное представление электрических колебаний. Допустим, что $z(t)$ — комплексная функция t , чей комплексный частотный спектр $s_z(f)$ равен нулю вне участка частот $(-W, W)$, как это предполагалось для определенных выше функций y , u и n . Хорошо известно¹⁾, что $z(t)$ может быть полностью

¹⁾См. часть I, теорема 13. (Прим. ред.)

описана при помощи ряда дискретных значений $z\left(\frac{r}{2W}\right)$ и, следовательно, представлена в виде комплексного многомерного вектора \mathbf{z} с этими составляющими.

Скалярное произведение двух таких векторов \mathbf{z} и \mathbf{w} равно

$$\mathbf{z}^*\mathbf{w} = \sum_r z_r^* w_r = 2W \int z^*(t)w(t) dt. \quad (6)$$

Второе уравнение может быть доказано путем последовательного применения теоремы Парсеваля¹⁾ к интегралам Фурье, представляющим функции $z(t)$ и $w(t)$, и к рядам Фурье, представляющим $s_z(f)$ и $s_w(f)$ в интервале $(-W, W)$. Из уравнения (6) немедленно следует

$$\mathbf{u}^*\mathbf{u} = |\mathbf{u}|^2 = 2W \int |u(t)|^2 dt = 4WE, \quad (7)$$

где E обозначает энергию колебаний u или ψ внутри участка интегрирования. Далее, если \mathbf{u}' описывает колебания $\frac{du}{dt}$, можно определить эффективную полосу частот β из уравнения

$$|\mathbf{u}'|^2 = \beta^2 |\mathbf{u}|^2. \quad (8)$$

Эта полоса частот β с точностью до постоянного множителя совпадает с эффективной полосой частот по определению Гэбора, который показал, что $\frac{\beta^2}{4\pi^2}$ есть момент второго порядка энергетического спектра u относительно начала отсчета, либо спектра ψ относительно f_0 .

Векторное обозначение особенно удобно при рассмотрении характеристик шумов, и результаты, полученные в этом разделе, будут использоваться в равной мере как в векторной, так и в интегральной форме, так как одна может быть преобразована в другую при помощи уравнения (6).

3. ОСНОВЫ ТЕОРИИ РАДИОЛОКАЦИОННОГО НАБЛЮДЕНИЯ

Центральной задачей является построение по данным принимаемым колебаниям и по известным переданным колебаниям распределения вероятностей для расстояния до цели, измеренного в единицах времени запаздывания τ . Допустим, что действительное значение времени запаздывания есть τ_0 , а коэффициента α — α_0 , тогда принимаемые колебания

$$\gamma(t) = \alpha_0 \psi(t - \tau_0) + v(t). \quad (9)$$

¹⁾ См. часть II, глава II, § 2.4. (Прим. ред.)

Чтобы построить распределение вероятностей для τ (в общем случае также для α), когда известны только $\gamma(t)$ и $\psi(t)$, необходимо прежде всего определить плотность вероятностей для наблюдаемых колебаний $\gamma(t)$ в предположении некоторых определенных значений τ и α . Эта плотность вероятностей зависит только от тех значений шумов, которые необходимо выбрать для удовлетворения принятых предположений.

При избранных значениях (τ , α) шумы определяются как

$$\gamma(t) - \alpha\psi(t - \tau)$$

или, пользуясь низкочастотными функциями (3):

$$[y(t) - \alpha u(t - \tau)e^{-2\pi if_0\tau}] e^{2\pi if_0 t} \quad (10)$$

Выражение в квадратных скобках может быть представлено вектором

$$\mathbf{y} - \alpha \mathbf{u}(\tau) e^{-i\omega\tau}, \quad (11)$$

где $\mathbf{u}(t - \tau)$ становится непрерывной векторной функцией от τ , а $\omega = 2\pi f_0$.

Вектор (11) однозначно определяет высокочастотные флуктуационные шумы, и в предположении шумов, подчиняющихся нормальному закону, его многомерная плотность вероятностей пропорциональна

$$\exp \left[- \frac{|\mathbf{y} - \alpha \mathbf{u}(\tau) e^{-i\omega\tau}|^2}{2N} \right]. \quad (12)$$

Здесь N — средняя физическая мощность шумов, равная среднему квадрату как вещественной, так и мнимой части шумовых флуктуаций в каждой точке отображения¹⁾. «Дискретны» шумов не имеют между собой корреляции, поскольку предполагается, что шумы *равномерно* распределены по участку высоких частот шириной $2W$.

Если априорное распределение вероятностей для τ и α обозначить $p_0(\tau, \alpha)$, то по принципу «обратной вероятности» их апостериорное распределение будет²⁾

$$p(\tau, \alpha) = p_0(\tau, \alpha) \exp \left[- \frac{|\mathbf{y} - \alpha \mathbf{u}(\tau) e^{-i\omega\tau}|^2}{2N} \right] \quad (13)$$

Это есть распределение частот истинных значений τ и α по множеству, в котором принимаемые колебания $\gamma(t)$ фиксированы.

¹⁾ Как и в главе I, так в дальнейшем будут называться точки, в которых берутся значения функции $z\left(\frac{r}{2W}\right)$, образующие составляющие многомерного вектора этой функции. Эти составляющие по примеру предыдущей главы будут именоваться «дискретными». (Прим. ред.)

²⁾ Подробное изложение теоремы «обратной вероятности», устанавливающей связь между апостериорным и априорным распределениями вероятностей, приведено в главе I, § 1. (Прим. ред.)

Вместе с тем априорное распределение есть соответствующая частота по множеству из всех принимаемых колебаний. Оно должно быть точно определено, прежде чем идти дальше.

Так как в этой работе рассматривается количество данных, несомых принимаемыми сигналами определенной энергии, то предполагается, что коэффициент α постоянен и, следовательно, его априорное распределение

$$p_0(\alpha) = \delta(\alpha - \alpha_0). \quad (14)$$

Далее предполагается, что априорное распределение вероятностей для τ равномерно по фиксированному интервалу T , меньшему или равному периоду повторения передаваемых колебаний. Такое предположение не только является простейшим, но и наименее связывающим, поскольку оно представляет состояние наименьшего предварительного знания между пределами распределения. Поэтому внутри интервала T

$$p_0(\tau, \alpha) = \frac{1}{T} \delta(\alpha - \alpha_0). \quad (15)$$

Из (13) следует, что апостериорное распределение для τ равно

$$p(\tau) = \int_0^{\infty} p(\tau, \alpha) d\alpha = \exp \left[- \frac{|y - \alpha_0 \mathbf{u}(\tau) e^{-i\omega\tau}|^2}{2N} \right] \quad (16)$$

внутри априорного интервала T и равно нулю вне его повсюду.

Это выражение будет теперь преобразовано к более удобному виду. Разлагая в ряд и опуская члены, не зависящие от τ , получим¹⁾

$$p(\tau) = \exp \left\{ \frac{\alpha_0}{N} \operatorname{Re} \mathbf{y}^* [\mathbf{u}(\tau) e^{-i\omega\tau}] \right\} \quad (17)$$

Из числа членов, которые исчезли, $|\mathbf{y}|^2$ зависит от τ_0 , но не зависит от τ , тогда как член $|\mathbf{u}(\tau)|^2$ пропорционален энергии сигнала за интервал наблюдения и не зависит от τ , поскольку он охватывает целое число периодов. Перейдем теперь к интегральной форме

$$p(\tau) = \exp \left[\frac{\alpha_0}{N_0} \operatorname{Re} \int y^*(t) u(t - \tau) e^{-i\omega\tau} dt \right] \quad (18)$$

где

$$N = 2WN_0, \quad (19)$$

а N_0 — средняя мощность шумов в единице полосы частот.

¹⁾ Здесь и в дальнейшем Re обозначает вещественную часть соответствующего выражения, а Im — мнимую часть. (Прим. ред.)

Из (3) и (9), подставляя y , найдем

$$p(\tau) = \exp \operatorname{Re} \left\{ \frac{1}{N_0} e^{-i\omega(\tau-\tau_0)} \int [\alpha_0^2 u^*(t-\tau_0)u(t-\tau) + \alpha_0 n^*(t)u(t-\tau) e^{-i\omega\tau_0}] dt \right\} \quad (20)$$

Так как $n(t)$ — комплексная случайная величина, постоянная $e^{-i\omega\tau_0}$ может быть включена в нее без изменения ее статистических свойств и поэтому в дальнейшем опущена. Обозначив фазу интеграла через θ и написав

$$g(\tau) = \frac{\alpha^2}{N_0} \int u^*(t-\tau_0)u(t-\tau) dt, \quad (21)$$

$$h(\tau) = \frac{\alpha_0}{N} \int n^*(t)u(t-\tau) dt, \quad (22)$$

получим (20) в виде

$$p(\tau) = \exp \{ \cos [\omega(\tau-\tau_0) - \theta] \cdot |g(\tau) + h(\tau)| \}, \quad (23)$$

Косинус — высокочастотная функция τ , тогда как θ и модуль — низкочастотные функции. Это означает, что апостериорное распределение состоит из последовательности близко расположенных пиков, лежащих под медленно изменяющейся огибающей. Эти пики возникают из сравнений, которые может делать наблюдатель между фазами несущих передаваемых и принимаемых колебаний.

Сведения, даваемые этой тонкой структурой, бесполезны, когда имеется неопределенность в различении одного пика от другого. Действительное распределение для τ , обозначаемое $P(\tau)$, можно найти путем сглаживания кривой $p(\tau)$ при устранении тонкой структуры. Важно, однако, иметь в виду, что хотя сведения, которые могли бы быть получены при сравнении фаз передаваемой и принимаемой несущих, не будут приниматься во внимание, в дальнейшем предполагается наличие фазовой когерентности от одного периода повторения модуляции до другого в течение интервала наблюдения. Другими словами, изменение фазы дает полезную информацию, но абсолютная величина — нет.

Сглаженное распределение легко получить интегрированием (23)

по τ в пределах от τ до $\tau + \frac{2\pi}{\omega}$, полагая [в этом интервале изменения θ и модуля незначительными. Это равноценно взятию огибающей $p(\tau)$, поскольку тонкая структура не имеет синусоидального характера. Интегрирование дает

$$P(\tau) = \lambda I_0(g + h), \quad (24)$$

где I_0 — видоизмененная функция Бесселя; λ — нормирующая постоянная. Следует подчеркнуть, что интегралы g и h распространяются на весь период наблюдения.

Свойства $P(\tau)$ в сильной степени зависят от относительной роли g и h , которые можно рассматривать как составляющие распределения вероятностей, связанные соответственно с сигналом и шумами. Для удобства они будут называться функцией сигнала и функцией шумов. Первая представляет собой функцию автокорреляции сигнала, а вторая — функцию взаимной корреляции между сигналом и шумами.

4. ФУНКЦИИ СИГНАЛА И ШУМОВ

В этом разделе будут сформулированы некоторые свойства $g(\tau)$ и $h(\tau)$ и попутно выяснен критерий оценки количества энергии, которое должны нести эхо-сигналы для удовлетворительного радиолокационного наблюдения.

Следует ожидать, что $P(\tau)$ должно иметь пик в окрестности истинной дальности τ_0 . Этот пик должен появиться из-за функции сигнала $g(\tau)$. Во всех других местах $g(\tau)$ будет практически незначительной вследствие специального выбора формы радиолокационных колебаний. Вблизи τ_0 можно разложить $u(t-\tau)$ по теореме Тейлора и получить из (21)

$$g(\tau) = \frac{\alpha_0^2}{N_0} \int u^*(t-\tau_0) \left[u(t-\tau_0) - (\tau-\tau_0) u'(t-\tau_0) + \frac{1}{2} (\tau-\tau_0)^2 u''(t-\tau_0) \right] dt, \quad (25)$$

пренебрегая высшими степенями $\tau-\tau_0$ и обозначая $\frac{du}{dt} = u'$. Положив

$$\rho^2 = \frac{2E}{N_0}, \quad (26)$$

где E — теперь энергия принимаемого сигнала $\alpha_0\psi$, можно получить из (7), (4) и (8) следующие полезные соотношения:

$$\frac{\alpha_0^2}{N_0} \int u^*(t-\tau_0) u(t-\tau_0) dt = \rho^2, \quad (27)$$

$$\frac{\alpha_0^2}{N_0} \int u^*(t-\tau_0) u'(t-\tau_0) dt = 0, \quad (28)$$

$$\frac{\alpha_0^2}{N_0} \int u^*(t-\tau_0) u''(t-\tau_0) dt = -\rho^2 \beta^2. \quad (29)$$

Это — определенные интегралы, которые берутся по целому числу периодов повторения функции $u(t)$. Второе соотношение зависит от определения (4) несущей частоты, а третье получается после интегрирования по частям. Поэтому¹⁾

$$g(\tau) = \rho^2 - \frac{1}{2} \rho^2 \beta^2 (\tau - \tau_0)^2 + O(\tau - \tau_0)^3. \quad (30)$$

Как будет видно, пик распределения $g(\tau)$ в точке τ_0 имеет высоту ρ^2 , причем в его окрестности $g(\tau)$ вещественна. Величина ρ^2 имеет важное значение, это — безразмерный энергетический параметр, постоянно встречающийся в излагаемой теории. Полученное выше разложение $g(\tau)$ понадобится в дальнейшем при определении точности измерения дальности.

Теперь необходимо рассмотреть функцию шумов $h(\tau)$ и ее влияние на $g(\tau)$. Преобразование (22) к векторному обозначению дает

$$h(\tau) = \frac{\alpha_0}{N} \mathbf{p}^* \mathbf{u}(\tau). \quad (31)$$

Это скалярное произведение является линейной функцией составляющих \mathbf{p} . Результат представляет собой стационарную случайную функцию, причем распределения действительной и мнимой частей независимы и являются нормальными относительно нуля. Пользуясь аддитивными свойствами дисперсий, легко показать, что

$$\overline{[\operatorname{Re} h(\tau)]^2} = \overline{[\operatorname{Im} h(\tau)]^2} = \frac{\alpha_0^2}{N} \overline{|\mathbf{u}(\tau)|^2} = \rho^2. \quad (32)$$

Далее, хорошо известно (часть II, глава III), что распределение амплитуд такой функции подчиняется закону

$$q(|h|) = \frac{1}{\rho^2} |h| e^{-\frac{|h|^2}{2\rho^2}} \quad (33)$$

и что значение среднего квадрата $|h|$ равно $2\rho^2$.

Когда $E < N_0$, эффективное значение $\rho\sqrt{2}$ модуля функции $h(\tau)$ больше пикового значения ρ^2 функции $g(\tau)$ и в $P(\tau)$ функции сигнала и шумов не различимы. Поэтому необходимый, хотя и недостаточный критерий удовлетворительного наблюдения заключается в том, что полная принимаемая энергия должна быть больше мощности шумов, приходящейся на единицу полосы частот (для удобства вычислений будем предполагать «много больше»).

¹⁾Знак O перед скобками указывает тот наибольший член разложения в ряд, который в дальнейшем отбрасывается. (Прим. ред.)

Условие

$$\rho^2 \gg 2 \text{ или } E \gg N_0 \quad (34)$$

будет применяться на протяжении дальнейшей части работы. Оно недостаточно, поскольку само по себе не гарантирует, что площадь под кривой $P(\tau)$ вся сосредоточена вблизи τ_0 .

Немедленным следствием этого условия является то, что пик функции сигнала $g(\tau)$ велик сравнительно с единицей.

Из (24) можно видеть, что это свойство еще усилено в $P(\tau)$, так как функция $I_0(x)$ асимптотически возрастает по показательному закону вместе с x . Поэтому для практических радиолокационных колебаний пик кривой нормированного распределения вероятностей, вызванный функцией сигнала, будет сконцентрирован вблизи τ_0 и в (30) высшими степенями $(\tau - \tau_0)$ можно будет пренебречь.

Исключая области вблизи этого пика, функция сигнала будет иметь незначительное влияние на $P(\tau)$, и функция шумов (если она достаточной величины) одна будет определять это распределение. Поэтому полный интервал T , в котором определена функция $P(\tau)$, можно разделить на два участка. Первый участок включает пик $P(\tau)$, созданный функцией сигнала, — здесь $P(\tau)$ будет обозначаться $P_g(\tau)$. Остальной участок имеет длительность, почти равную T , где $P(\tau)$ будет обозначаться $P_h(\tau)$, так как функция g почти не оказывает здесь влияния. Можно удовлетворительно разделить эти два участка, потому что $P_h(\tau)$ само состоит (когда ρ^2 велико) из последовательности изолированных и случайно возникающих пиков.

Первый участок будет определять точность, с которой может быть измерена дальность, тогда как площадь под вторым участком определяет степень ненадежности наблюдения.

Под ненадежностью наблюдения понимается вероятность того, что во множестве результатов наблюдений при фиксированных принимаемых колебаниях действительное значение дальности не лежит где-либо вблизи τ_0 .

Теперь наиболее важные свойства функций сигнала и шумов уже изложены и, как видно, если полная принимаемая энергия невелика по сравнению с мощностью шумов, приходящейся на единицу полосы частот, то от системы может быть получено только незначительное количество данных. Этот результат относится, конечно, не только к радиолокации; он следует в весьма общем случае из уравнения (1).

5. ТОЧНОСТЬ ИЗМЕРЕНИЯ ДАЛЬНОСТИ

Могут быть два различных статистических подхода к проблеме об ошибке наблюдения. Первый подход (обманчиво простой) заключается в фиксировании истинного значения и определении

разброса оценок. Другой, применяемый в этой работе, состоит в фиксировании типичной формы принимаемых колебаний, отраженных от цели, расположенной на истинном расстоянии τ_0 , и в рассмотрении разброса истинных значений τ , которые могли бы быть созданы такими принимаемыми колебаниями. Последний подход, связанный с введением распределения вероятностей $P(\tau)$, устраняет рассмотрение того, как оператор оценивает истинное значение, когда наблюдает принимаемые колебания.

Так как распределение $P(\tau)$ представляет свойства множества, в котором принимаемые колебания γ фиксированы, оно является *частным*, поскольку зависит от данного выбора γ . Поэтому необходимо рассмотреть средние статистические свойства $P(\tau)$ путем изменения γ , сохраняя τ_0 постоянным [см. уравнение (9)]. Средние значения по этому множеству при фиксированном τ_0 будут различаться от средних значений по множеству при фиксированном γ путем применения символа C_p вместо черты. Дальше будет показано, что распределение $P_g(\tau)$ является приблизительно нормальным вблизи τ_0 , и вычислено его стандартное отклонение.

До сих пор любая ненадежность наблюдения, возникающая из-за поведения функции шумов $h(\tau)$ вне g -области вокруг τ_0 , игнорировалась, но ее влияние внутри этой области представляет интерес и может быть учтено следующим образом. Из уравнения (31) и разложения в ряд Тейлора $u(l-\tau)$ функция шумов вблизи τ_0 равна

$$[h(\tau) = \frac{\alpha_0}{N} [n^*u(\tau_0) - (\tau - \tau_0)n^*u'(\tau_0) + C(\tau - \tau_0)^2]. \quad (35)$$

Можно показать, что, если r^2 достаточно велико, пренебрежение всеми членами, за исключением линейного и постоянного членов, не является серьезным, несмотря на необходимость включения квадратичного члена в соответствующее разложение $g(\tau)$. Поэтому можно написать

$$h(\tau) = \xi + i\eta = [\xi_0 + (\tau - \tau_0)\xi'] + i[\eta_0 + (\tau - \tau_0)\eta'], \quad (36)$$

где ξ_0, η_0, ξ' и η' — случайные величины во множестве с фиксированным τ_0 , обладающие нормальным распределением вокруг нуля. Из уравнения (32)

$$C_p(\xi_0^2) = C_p(\eta_0^2) = r^2, \quad (37)$$

и при помощи таких же аргументов легко показать, что

$$C_p(\xi'^2) = C_p(\eta'^2) = r^2 r^2. \quad (38)$$

Таким образом, $|g+h|$ может быть приблизительно оценено в окрестности τ_0 для последующей подстановки в уравнение (24).

Разлагая модули в ряды и пренебрегая всеми степенями $(\tau - \tau_0)$ выше второй, получаем из уравнений (30) и (36)

$$|g(\tau) + h(\tau)| = \rho^2 + \xi_0 + \frac{\eta_0^2}{2\rho^2} + \frac{\xi'^2}{2\rho^2\beta^2} - \frac{1}{2} \rho^2 \beta^2 \left(\tau - \tau_0 + \frac{\xi'}{\beta^2 \rho^2} \right)^2 + O\left(\frac{1}{\rho^2}\right). \quad (39)$$

Это выражение можно записать в более простой форме

$$|g + h| \approx \rho^2 + \chi - \frac{1}{2} \rho^2 \beta^2 (\tau - \tau_m)^2, \quad (40)$$

где из (37) и (38)

$$\text{Cp } \chi = 1, \quad (41)$$

$$\text{Cp } \tau_m = \tau_0, \quad (42)$$

$$\text{Cp } (\tau_m - \tau_0)^2 = \frac{1}{\rho^2 \beta^2}. \quad (43)$$

Поэтому из уравнения (24) имеем вблизи τ_0

$$P_g(\tau) = \lambda I_0 \left[\rho^2 + \chi - \frac{1}{2} \rho^2 \beta^2 (\tau - \tau_m)^2 \right]. \quad (44)$$

Асимптотическое разложение I_0 равно

$$I_0(x) \approx \frac{e^x}{\sqrt{2\pi x}} \left[1 + \frac{1}{8x} + O\left(\frac{1}{x^2}\right) \right]. \quad (45)$$

Тривиальное применение этого разложения показывает, что распределение $P_g(\tau)$ должно быть вблизи τ_0 приближенно нормальным с параметрами

$$\tau = \tau_m, \quad (46)$$

$$\overline{(\tau - \tau_m)^2} = \frac{1}{\beta^2 \rho^2}. \quad (47)$$

Поэтому стандартное отклонение τ от его среднего значения в $P_g(\tau)$

$$\delta\tau = \frac{1}{\rho\beta} \quad (48)$$

Это выражение представляет апостериорную неопределенность измерения дальности. То обстоятельство, что среднее значение τ в $P_g(\tau)$ само имеет такую же неопределенность по отношению к τ_0 [уравнение (43)], не является простым совпадением. Именно этот результат следовало ожидать в результате применения «обратной вероятности»: оба разброса значений в действительности представляют один и тот же эффект.

Уравнение (48) представляет первый результат, имеющий практический интерес. Неопределенность не зависит от частотной характеристики приемника, поскольку она была найдена безотносительно от какого-то определенного приемного устройства и представляет оптимальные свойства системы.

6. НЕНАДЕЖНОСТЬ НАБЛЮДЕНИЯ

В конце § 4 было отмечено, что степень ненадежности наблюдения есть вероятность того, что во множестве результатов наблюдений при фиксированном принимаемом сигнале действительная дальность τ не находится где-то вблизи τ_0 . Поэтому ненадежность A может быть представлена площадью, лежащей под той частью кривой $P(\tau)$, которая связана только с функцией шумов $h(\tau)$, т. е. в которой исключен пик сигнала, рассмотренный в предыдущем разделе. Она будет вычислена умножением ожидаемой площади, рассчитанной на единицу дальности τ и связанной только с одной функцией h , на полный априорный интервал T . Последний включает небольшой участок интегрирования вблизи τ_0 , который, строго говоря, следовало бы исключить, но так как для любой практической системы справедливо неравенство

$$T \gg \frac{1}{\rho\beta}, \quad (49)$$

то это приближение является оправданным.

Удобно определить $P_h(\tau)$ в виде

$$P_h(\tau) = \lambda I_0(|h|), \quad (50)$$

что совпадает с видом выражения (24) для $P(\tau)$, за исключением области вблизи τ_0 . «Ожидаемая площадь на единицу дальности» под кривой $P_h(\tau)$ является, конечно, ожидаемым значением P_h и может быть получена рассмотрением распределения вероятностей для $|h|$ (уравнение 33). Поэтому

$$\text{Ср } P_h = P_h = \int_0^{\infty} P_h Q(P_h) dP_h, \quad (51)$$

где

$$Q(P_h) dP_h = q(|h|) d|h|. \quad (52)$$

Оба средних значения в (51) равны, поскольку $h(\tau)$ стационарная случайная функция. Интеграл может быть вычислен подстановкой q из (33) и P_h из (50), пользуясь разложением (45) для функции Бесселя. В результате получим

$$\bar{P}_h = \lambda e^{\rho^2/2} [1 + O(\rho^{-4})], \quad (53)$$

и ненадежность, определяемая средней площадью под кривой $P_h(\tau)$, равна

$$A = T\bar{P}_h. \quad (54)$$

Остается только определить нормирующий коэффициент λ для $P(\cdot)$, зависящий также от площади под кривой $P_g(\tau)$, связанной с функцией сигнала.

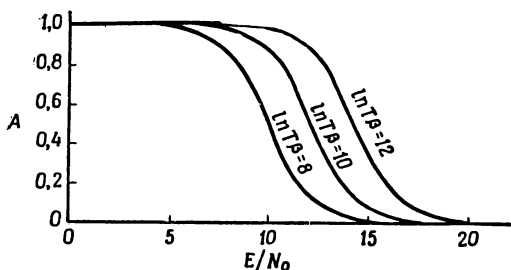
Эта площадь получается интегрированием распределения [уравнение (44)], пользуясь опять разложением функции Бесселя.

В результате находим¹⁾

$$\int_T P_g(\tau) d\tau = \frac{\lambda e^{\rho^2+1}}{\rho^{2\beta}} \left[1 - \frac{1}{8\rho^2} + O(\rho^{-4}) \right]. \quad (55)$$

Пренебрегая по сравнению с единицей членами порядка ρ^{-4} , получаем нормирующее уравнение

$$\begin{aligned} 1 &= \int_T P_g(\tau) d\tau + \int_T P_h(\tau) d\tau = \\ &= \lambda \left[\frac{e^{\rho^2+1}}{\rho^{2\beta}} \left(1 - \frac{1}{8\rho^2} \right) + T e^{\rho^2/2} \right]. \end{aligned} \quad (56)$$



Фиг. 1. Порог разборчивости.

Отсюда приближенно

$$A = \frac{T\rho^{2\beta}}{T\rho^{2\beta} + e^{\frac{1}{2}\rho^2+1}}. \quad (57)$$

Таким образом, при постоянных T и β ненадежность уменьшается от значения, близкого к единице, до нуля по мере возрастания $\frac{1}{2}\rho^2$ —

отношения энергии принимаемого сигнала к мощности шумов, приходящейся на единицу полосы частот. Как видно из фиг. 1, изменение ненадежности происходит довольно резко, благодаря чему образуется порог разборчивости. Условно его можно определить значением $A=0,5$, и, следовательно, при этом

$$\frac{1}{2}\rho^2 + 1 = \lg T\rho^{2\beta}. \quad (58)$$

¹⁾ Здесь ради удобства использовано среднее значение показательной функции [см. уравнение (41)]. Для дальнейшего это не имеет существенного значения, за исключением $A \approx 0,5$, когда в любом случае неизбежны дополнительные приближения.

При $\left(\frac{1}{2} \rho^2 + 1\right) \gg \log T \rho^2 \beta$ обеспечивается надежный прием. Это второй критерий успешного радиолокационного наблюдения.

7. КОЛИЧЕСТВО ДАННЫХ ПРИ РАДИОЛОКАЦИОННОМ НАБЛЮДЕНИИ

Количество данных, получаемых при радиолокационном наблюдении, равно разности «энтропий»¹⁾ априорного и апостериорного распределений вероятностей для дальности. Поскольку было предположено, что априорное распределение — прямоугольное с шириной T , его «энтропия», определяемая в виде²⁾

$$H = - \int p \log p \, d\tau, \quad (59)$$

равна

$$H_0 = \log T \quad (60)$$

Искомая «энтропия» апостериорного распределения есть средняя «энтропия» для $P(\tau)$, взятая по полному множеству возможных принимаемых колебаний. Поскольку, однако, «энтропия» не зависит от τ_0 , достаточно рассмотреть среднее по множеству при фиксированном τ_0 . Как и раньше, часть распределения вблизи τ_0 , а именно $P_g(\tau)$, и остающаяся часть, $P_h(\tau)$, будут рассматриваться раздельно, причем полная «энтропия» равна сумме.

Возникает, однако, трудность при вычислении «энтропии» некоторого данного распределения $P(\tau)$, поскольку при этом необходимо использовать правильное значение нормирующего коэффициента данного распределения, который изменяется по множеству с фиксированным τ_0 . Благодаря этому трудно выполнить интегрирование и для упрощения используется постоянное значение нормирующего коэффициента λ , зависящее от средних свойств $P(\tau)$ и определяемое уравнением (56).

Флуктуации λ при $\rho^2 \gg 1$ связаны почти исключительно с $P_h(\tau)$, и пренебрежение ими оправдывается в любом из следующих двух случаев:

а) если полная площадь под кривой $P_h(\tau)$ мала, т. е. $A \approx 0$, и «энтропия» приближенно равна «энтропии» одного только распределения $P_g(\tau)$;

б) если полный интервал T , по которому берется $P_h(\tau)$, настолько велик, что P_h включает достаточный статистический образец тех свойств множества, которые влияют на «энтропию».

Это условие приводит к $A \approx 1$, хотя вывод длинен и здесь опускается. Тогда «энтропия» связана почти полностью с $P_h(\tau)$.

¹⁾См. часть I, § 11. (Прим. ред.)

²⁾Здесь $p(\tau)$ означает некоторое распределение вероятностей.

Таким образом, сначала следует вычислить среднюю «энтропию», связанную только с одним $P_g(\tau)$, которая даст «энтропию» распределения $P(\tau)$ для $A \approx 0$. Затем нужно определить «энтропию» за счет одного только $P_h(\tau)$, которая представляет «энтропию» распределения $P(\tau)$ для $A \approx 1$. Наконец, может быть найдено приближение для промежуточной (пороговой) области при пренебрежении флуктуациями нормирующего коэффициента.

Часть «энтропии», связанная с распределением $P_g(\tau)$, может быть найдена подстановкой в уравнение (59) усредненного выражения для $P_g(\tau)$ из уравнений (44) и (41), пользуясь разложением I_0 и выполняя интегрирование. Это дает

$$\text{Ср } H_g = - \frac{\lambda e^{\rho^2+1}}{\rho^2\beta} \left[\left(1 - \frac{1}{8\rho^2} \right) \log \frac{\lambda}{\rho\sqrt{2\pi}} + \rho^2 + \frac{3}{8} + O(\rho^{-2}) \right]. \quad (61)$$

Но если $A \approx 0$, то из (55) нормирующий коэффициент определяется соотношением

$$\frac{\lambda e^{\rho^2+1}}{\rho^2\beta} \left(1 - \frac{1}{8\rho^2} \right) = 1, \quad (62)$$

и, таким образом:

$$\text{Ср } H_g = \log \frac{\sqrt{2\pi e}}{\rho\beta} + O(\rho^{-2}). \quad (63)$$

Количество данных равно разности априорной и апостериорной «энтропий», т. е.

$$I = H_0 - H_1 = \log T - H_1 \quad (64)$$

и, следовательно:

$$I \approx \log \frac{T\rho\beta}{\sqrt{2\pi e}}, \quad (A \approx 0). \quad (65)$$

Поскольку $P_h(\tau)$ — стационарная случайная функция, то можно вычислить среднее значение связанной с ней «энтропии» умножением ожидаемого значения $P_h \log P_h$ на интервал интегрирования T . Тогда, пренебрегая флуктуациями λ :

$$\text{Ср } H_h = - T \int_0^{\infty} Q(P_h) P_h \log P_h dP_h. \quad (66)$$

Отсюда, подставляя значения $Q(P_h)$ и P_h из уравнений (52), (33) и (50), пользуясь разложением I_0 и интегрируя, получим

$$\text{Ср } H_h = - \lambda T e^{\rho^2/2} \left[\log \frac{\lambda}{\rho\sqrt{2\pi}} + \rho^2 + \frac{1}{2} + O(\rho^{-2}) \right] \quad (67)$$

Но если $A \approx 1$, то, пользуясь (53) и (54), найдем приближенное значение нормирующего коэффициента λ из

$$\lambda e^{\rho^2/2} = \frac{1}{T} \quad (68)$$

Тогда

$$\text{Ср } H_h = \log(T\rho\sqrt{2\pi}) - \frac{1}{2}\rho^2 - \frac{1}{2} + O(\rho^{-2}) \quad (69)$$

и, следовательно:

$$I \approx \frac{1}{2}\rho^2 - \log(\rho\sqrt{2\pi}) + \frac{1}{2}, \quad (A \approx 1). \quad (70)$$

Уравнения (65) и (70) дают приближенные значения количества данных при условии надежного и ненадежного приема за исключением области вблизи порога при $A=0,5$.

Когда A не лежит вблизи нуля или единицы, необходимо, конечно, нормировать $P(\tau)$, приняв во внимание как $P_g(\tau)$, так и $P_h(\tau)$, и пренебрегая флуктуациями нормирующего коэффициента, которые должны быть в этой области учтены. Значение λ из уравнения (56) подставляется в уравнения (61) и (67), сложение которых дает приближенное среднее значение апостериорной «энтропии» H_1 . После вычитания из H_0 получим

$$I \approx -\log\left[\sqrt{2\pi e}\left(\rho e^{-\frac{\rho^2}{2}-1} + -\frac{1}{T\rho\beta}\right)\right] + O(\rho^{-2}). \quad (71)$$

Как будет видно, в зависимости от того, преобладает ли под знаком логарифма первый или второй член, т. е. $A \approx 0$ или $A \approx 1$, это уравнение сводится, как это и должно быть, к (65) или (70). Оно представляет лучшее приближение для полного количества данных, чем другие ранее приведенные уравнения, несмотря на пренебрежение флуктуациями нормирующего коэффициента.

8. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Основные результаты изложенной теории иллюстрирует фиг. 2, где представлены зависимости количества данных I [уравнение (71)] и ненадежности A [уравнение (57)] от величин $\ln T\beta$ и $\frac{E}{N_0}$.

Как видно, линии I имеют совершенно разный характер по обе стороны от порога разборчивости ($A=0,5$). Когда ненадежность A мала, линии I асимптотически приближаются к пунктирным кривым, даваемым уравнением (65). Получаемые сведения в этой области обеспечивают почти надежное знание дальности, подверженной «топологической» ошибке, которая по § 5 имеет стандартное отклонение

$$\delta\tau = \frac{1}{\rho\beta}$$

где

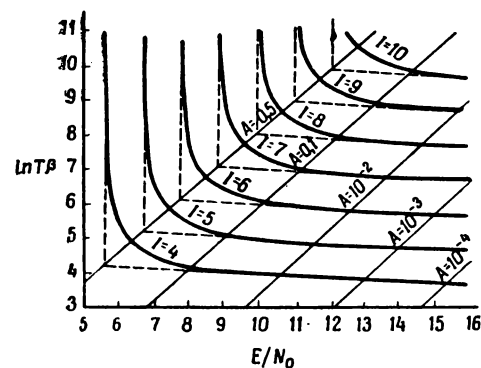
$$\rho = \sqrt{\frac{2E}{N_0}}$$

а β — «эффективная полоса частот» передаваемых колебаний.

Как и следует ожидать, ошибка уменьшается по мере возрастания полной принимаемой энергии E , если мощность шумов на единицу полосы пропускания N_0 остается неизменной. Она также уменьшается при увеличении β , однако увеличение одной полосы

частот β с целью получения более высокой разрешающей способности имеет свои недостатки, поскольку наблюдение постепенно становится более ненадежным.

Когда ненадежность больше половины, графики I асимптотически приближаются к пунктирным вертикальным линиям, даваемым уравнением (70), которые представляют верхний предел количества данных при постоянной принимаемой энергии.



Фиг. 2. Кривые количества данных и ненадежности.

В общей теории (часть I) имеется подобный же верхний предел, с которым можно провести сравнение. В весьма общем случае количество данных I , которое может быть извлечено из некоторого принимаемого сообщения с энергией E , ограничено величиной $\frac{E}{N_0}$ натуральных единиц. Этот результат получается из (1), если написать

$$I = Ct = Wt \log \left(1 + \frac{E}{N_0 Wt} \right) \quad [\text{нат. ед.}],$$

где t — длительность сообщения. При $Wt \rightarrow \infty$ I достигает предельного значения

$$I_s = \frac{E}{N_0} = \frac{1}{2} \rho^2.$$

Соответствующий предел в радиолокационном случае при $\beta \rightarrow \infty$ меньше этого на $\left(\log \rho \sqrt{2\pi} - \frac{1}{2} \right)$ натуральных единиц. Можно показать, что расхождение вызывается данными, связанными с тонкой структурой, которые были умышленно игнорированы при переходе от уравнения (23) к (24). Если бы в радиолокации можно

было работать без высокочастотной несущей, то этого явления не возникало бы и предельные значения в ненадежной области (фиг. 2) точно соответствовали бы теоретическому пределу.

При нормальном приеме величины β , N_0 и априорный интервал T постоянны, тогда как E возрастает вместе с временем наблюдения. Поэтому при рассмотрении диаграммы начинают слева и постепенно двигаются вдоль горизонтали направо. Вначале количество данных быстро возрастает при постепенном уменьшении ненадежности. Однако после пересечения порога и перехода в рабочую область количество данных возрастает очень медленно и представляет постепенное улучшение точности.

Это явление находится в полном соответствии с идеями статистической теории передачи электрических сигналов. Как только приближенное положение цели известно с надежностью, поступающая в дальнейшем энергия постепенно увеличивает количество данных, которые уже частично известны. Разрешающая способность по дальности возрастает за счет непрерывного уточнения, а не за счет систематического исключения ошибки. По этой причине наиболее эффективная радиолокационная система должна работать как можно ближе к порогу разборчивости (насколько близко, зависит от допустимой степени ненадежности).

Необходимо подчеркнуть, что в этой работе рассматриваются фазово-когерентные принимаемые колебания и это существенно ограничивает практическое применение теории. Если параметр дальности τ был бы сам функцией времени, то вводимые в теории интегралы по времени должны были бы вычисляться иначе. При этом полный период наблюдения должен быть ограничен интервалом, достаточно малым, чтобы в пределах его τ могла бы рассматриваться как постоянная. В этих условиях нет основания отдать предпочтение (за исключением практических соображений) концентрации используемой энергии в одном периоде модуляции или распределении ее по нескольким периодам. Нет сомнения, однако, что в последнем случае энергия тратится напрасно, если нельзя использовать длительное фазово-когерентное интегрирование.

В полученные формулы не входит полоса пропускания приемника. До тех пор, пока эта полоса достаточно широка, чтобы включить весь передаваемый спектр, количество данных не изменяется при ее изменении. Выбор формы частотной характеристики зависит только от наилучшего способа выделения данных, несомых принимаемыми колебаниями. Эта практическая задача в работе не рассматривается.

Г л а в а III

ПРОЕКТИРОВАНИЕ РАДИОЛОКАЦИОННЫХ ПРИЕМНИКОВ НА ОСНОВЕ СТАТИСТИЧЕСКИХ МЕТОДОВ ¹⁾

Ф. ВУДВОРД

1. ВВЕДЕНИЕ

Задача этой работы — изложить математический метод (в форме несколько идеализированного примера), при помощи которого в принципе всегда могут быть определены свойства теоретически идеального радиолокационного приемника.

До последнего времени общепринято рассматривать отношение сигнал/шум как основной параметр при проектировании радиолокационных приемников. Принимались меры для получения максимально возможного отношения сигнал/шум на выходе приемника. Такой подход представляется в настоящее время ошибочным, поскольку это отношение не служит мерой фактически получаемых при радиолокационном наблюдении данных. Путем пропускания колебаний через нелинейное устройство отношение сигнал/шум часто может быть искусственно увеличено, не изменяя вовсе количества данных.

В радиолокации необходимо дать ответ на такие вопросы, как наличие или отсутствие цели, ее дальность, движется ли она и т. д.

Попытка сконструировать приемник, дающий точные ответы на один или все эти вопросы, оказывается неудачной, так как показания неизбежно искажаются шумами. Но требование автоматического определения относительных вероятностей всех возможных ответов представляется вполне реальным, и никакое приемное устройство не может, вероятно, дать что-либо большее.

В дальнейшем будет показано, как эта идея используется для решения часто встречающейся задачи — определения времени запаздывания периодического сигнала известной формы и амплитуды. В радиолокации это означает измерение дальности до неподвижной цели, наличие которой и сила отраженного ею сигнала известны. Хотя эта задача и имеет искусственный характер, но она достаточна для иллюстрации предлагаемого метода и сама по себе не лишена практического интереса.

Количество данных, несомых радиолокационным сигналом, было вычислено в другой работе ²⁾ без указания какого-либо действи-

¹⁾ Ph. M. Woodward, «Information Theory and the Design of Radar Receivers», *Proc. of the IRE*, 39, № 12, 1521—1524, December 1951.

²⁾ См. главу II, часть III. (Прим. ред.)

тельного способа их получения. В настоящей работе делается упор на вопросы проектирования радиолокационного приемника, а не на оценку его действительных качественных показателей.

2. ПРИНЦИП «ОБРАТНОЙ ВЕРОЯТНОСТИ» И ФУНКЦИЯ КОЛИЧЕСТВА ДАННЫХ

Излагаемый метод целиком основан на принципе «обратной вероятности»¹⁾, устанавливающим связь между апостериорными и априорными вероятностями гипотез

$$P(H_n | Ob) = P(H_n) P(Ob | H_n). \quad (1)$$

Здесь $P(H_n | Ob)$ — апостериорная вероятность гипотезы H_n , после того как результат наблюдения станет известным; $P(H_n)$ — априорная вероятность гипотезы H_n ; $P(Ob | H_n)$ — вероятность наблюдения, если гипотеза H_n правильна. В рассматриваемом случае «гипотезами» являются все возможные временные запаздывания данного периодического сигнала. «Наблюдение» — просто данный принятый сигнал вместе с воздействующими на него шумами, запаздывающий на неизвестный промежуток времени. Предполагается, что шумы обладают равномерным спектром и подчиняются нормальному закону распределения вероятностей.

Следует иметь в виду, что «принятый сигнал» — это не сигнал на выходе приемника, поскольку в этом случае предreshался бы весь вопрос. Речь идет о сигнале на входе приемника, включая и шумы, которые могут быть созданы в самом приемнике.

Априорные вероятности гипотез образуют непрерывное распределение вероятностей для неизвестного временного запаздывания τ . Ради простоты это распределение принимается равномерным в промежутке времени, равном одному периоду сигнала. Впрочем, в излагаемую теорию могут быть введены и другие сведения, основанные на предварительном знании.

Наибольшим, что можно определить при помощи приемника на основании принятого сигнала, является апостериорное распределение вероятностей для временного запаздывания τ . Это распределение и является искомыми данными. Наиболее удобно иметь дело с его логарифмом $Q(\tau)$, который в дальнейшем будет называться «функцией количества данных».

Тогда уравнение (1) можно написать в виде

$$Q(\tau) = \log P(Ob | \tau) + \text{const}, \quad (2)$$

где $P(Ob | \tau)$ — плотность распределения вероятностей для принятого сигнала в случае гипотезы τ . Постоянный член есть просто логарифм нормирующего коэффициента апостериорного распределения; в дальнейшем он опущен.

¹⁾ Подробнее см. § 1, глава I, часть III. (Прим. ред.)

3. ВЫЧИСЛЕНИЕ ФУНКЦИИ КОЛИЧЕСТВА ДАННЫХ

Рассмотрим случай, когда истинное значение временного запаздывания сигнала есть τ_0 , и запишем принимаемый сигнал при помощи вещественных функций в виде

$$Y(t) = G(t - \tau_0) + I(t). \quad (3)$$

Здесь $G(t)$ — высокочастотный сигнал, который был бы принят при отсутствии шумов и временного запаздывания; предполагается, что он известен заранее. Допустим, что $G(t)$ периодически модулирован и хотя удобнее рассматривать импульсный сигнал, теория справедлива для любой периодической модуляции, включая частотную.

Функция $I(t)$ представляет воздействующие на сигнал шумы. Плотность распределения вероятностей величины $I(t)$ в любой данный момент времени t предполагается подчиняющейся нормальному закону, но это понятие необходимо обобщить. Можно показать, например, применением теоремы о разложении непрерывной функции¹⁾, либо методами статистической механики, что плотность вероятностей функции шумов $I(t)$ в целом, обладающая соответствующим числом измерений²⁾, пропорциональна

$$\exp \left[-\frac{1}{N_0} \int I(t)^2 dt \right] \quad (4)$$

где N_0 — средняя мощность шумов в единице полосы частот.

Для наблюдения доступен сигнал $Y(t)$, а не непосредственно временное запаздывание τ_0 , поэтому необходимо поочередно испытать все возможные значения τ . При избранной гипотезе τ можно заключить, что одни шумы должны описываться выражением $Y(t) - G(t - \tau)$, для которого плотность вероятностей пропорциональна

$$\exp \left\{ -\frac{1}{N_0} \int [Y(t) - G(t - \tau)]^2 dt \right\} \quad (5)$$

Следовательно, согласно уравнению (2):

$$Q(\tau) = -\frac{1}{N_0} \int [Y(t) - G(t - \tau)]^2 dt. \quad (6)$$

Таким образом, функция данных пропорциональна интегралу квадрата отклонения принимаемого сигнала от гипотетического, свободного от воздействия шумов, сигнала с временным запаздыванием τ . При изменении гипотетического временного запаздывания τ его значение, обеспечивающее минимум среднеквадра-

¹⁾ См. часть I, теорема 13. (Прим. ред.)

²⁾ Подробнее см. § 3, глава I, часть III. (Прим. ред.)

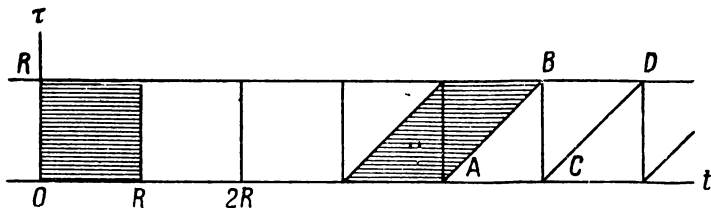
точного уклонения от принятого сигнала, соответствует максимуму функции количества данных. С точки зрения наблюдателя именно это значение τ есть наиболее вероятная величина временного запаздывания.

Пределы интегрирования в уравнении (6) выбираются в соответствии с исследуемой частью принятого сигнала; необходимо, чтобы она была равна целому числу периодов повторения модуляции. Если разложить подинтегральное выражение на три члена, то оказывается, что интеграл с G^2 не зависит от τ вследствие периодичности, а интеграл с Y^2 по уравнению (3) зависит от τ_0 , но не от τ . Следовательно, эти два члена можно исключить из выражения для $Q(\tau)$, так как все величины, не зависящие от τ , можно включить в ранее опущенный нормирующий множитель. Тогда остается

$$Q(\tau) = \frac{2}{N_0} \int Y(t)G(t - \tau)dt. \quad (7)$$

Можно сказать, что подинтегральное выражение существует для всех значений t , но τ сосредоточено внутри некоторого промежутка с определенными, заранее фиксированными, пределами, скажем, между 0 и R — периодом повторения сигнала $G(t)$. Область существования подинтегрального выражения может быть представлена схематически в виде полосы бесконечной длины в направлении t , шириной R в направлении τ (фиг. 1).

Если необходимо образовать функцию количества данных $Q(\tau)$, чтобы она представляла все данные, содержащиеся в принятом сигнале $Y(t)$ в интервале $(0, R)$, то $Y(t)G(t - \tau)$ должно быть



Фиг. 1. Два метода интегрирования данных о дальности цели.

проинтегрировано по t в пределах от 0 до R для каждого значения τ , как это показано на фиг. 1 заштрихованным квадратом. Последующие наблюдения отраженных сигналов связаны, очевидно, с дальнейшими этапами интегрирования, пределы которых всегда будут отмечаться на диаграмме вертикальными линиями.

Неудобство этого процесса с практической точки зрения состоит в том, что все данные, накопленные за один какой-то интервал, становятся доступными мгновенно в конце интервала, а затем следует перерыв, пока выполняется новое интегрирование.

Существует другой более естественный, хотя математически

и менее прямой метод. При выполнении интегрирования не в последовательных квадратах, а в параллелограммах, как показано на диаграмме, функция количества данных становится известной постепенно во времени от $\tau=0$ до $\tau=R$ (от A до B на графике). В тот момент времени, когда достигнута точка B , начинается новый цикл от C до D . Таким образом, устанавливается зависимость между τ и t , имеющая как бы форму пилообразной кривой временной развертки.

Этот процесс «прогрессивного» интегрирования соответствует, конечно, не фиксированным пределам в интеграле уравнения (7), а пределам, совместно перемещающимся во времени. Действительно, значение функции количества данных в момент времени $t = nR + \tau$, где n относится к n -му циклу развертки, а τ заключено между 0 и R , равно

$$Q_n(\tau) = \frac{2}{N_0} \int_{t-R}^t Y(t)G(t-\tau) dt. \quad (8)$$

С точки зрения математического исследования неудобно, что функция количества данных есть логарифм распределения апостериорных вероятностей для τ только в том случае, когда пределы интегрирования постоянны. Другими словами, каждое гипотетическое значение τ должно быть испытано, строго говоря, на одной и той же части принятого сигнала. Здесь нет возможности подробно рассмотреть этот вопрос, следует только указать, что «прогрессивное» распределение вероятностей

$$P_n(\tau) = e^{Q_n(\tau)} \quad (9)$$

(нормирующий множитель опущен) достаточно для всех практических целей как точное апостериорное распределение. В частности, если данные, полученные в последовательные периоды принятого сигнала, объединяются либо путем суммирования Q_n по n , либо перемножением P_n , результирующее распределение отличается от истинного апостериорного распределения только вследствие краевых эффектов, которые постепенно уменьшаются.

4. ИДЕАЛЬНЫЙ ФИЛЬТР, ВКЛЮЧЕННЫЙ ПЕРЕД ДЕТЕКТОРОМ

Функция количества данных, определяемая уравнением (8), допускает очень простую интерпретацию. Форма этого выражения, представляющего линейную суперпозицию, аналогична выражению для сигнала на выходе линейного фильтра. Действительно, это — сигнал на выходе фильтра в момент времени $t = nR + \tau$ при подаче на вход принимаемого сигнала $Y(t)$. При этом импульсивная реак-

ция фильтра определяется как

$$\delta(t) \rightarrow \begin{cases} \frac{2}{N_0} G(t - \tau), & 0 < t < R, \\ 0, & t < 0 \text{ и } t > R. \end{cases} \quad (10)$$

Подобный фильтр (если не считать коэффициента пропорциональности $\frac{2}{N_0}$) был рассмотрен Ван-Влеком и Миддльтоном¹⁾, показавшими, что он является единственным линейным фильтром, обеспечивающим максимум отношения импульсного сигнала к шумам. (Это свойство не имеет, однако, значения для излагаемой теории.) Амплитудно-частотная характеристика фильтра имеет такую же форму, что и амплитудный спектр одного периода входного сигнала G , а его фазово-частотная характеристика по абсолютным значениям равна, а по знаку противоположна фазовому спектру G .

Конечно, на выходе фильтра получается модулированный высокочастотный сигнал. Если его пропустить через выпрямитель с показательной характеристикой, соответствующей уравнению (9), то он будет описываться выражением, аналогичным уравнению «прогрессивного» распределения апостериорных вероятностей для τ . Выше уже указывалось, что если данные, полученные в отдельные периоды, должны быть объединены, то перед выпрямлением несколько циклов Q может быть сложено вместе. Эта операция представляет собой просто фазово-когерентное сложение, производимое от импульса к импульсу. Его следует делать только в том случае, когда τ совершенно не зависит от времени, как до сих пор предполагалось.

Действие этого идеального фильтра сводится к тому, что выходной импульсный сигнал становится сходным с особенно большим выбросом шумов; все структурные особенности, отличавшие первоначально сигнал от шумов, устранены и преобразованы в разницу по амплитуде. Это может показаться странным, поскольку сигнал и шумы на выходе имеют разные спектры мощности. К тому же, если сигнал на входе представляет собой прямоугольный импульс, шумы на выходе можно рассматривать как множество перекрывающихся прямоугольных импульсов, тогда как сигнал на выходе будет треугольным импульсом. Остается, однако, фактом, что множество прямоугольных импульсов, перекрывающихся на микроскопических интервалах при образовании шумов, подчиняющихся нормальному закону, создает такой фон, на котором невозможно различить очертания одиночного треугольного импульса. Если бы действительно особенности структуры выходно-

¹⁾ Дж. Ван-Влек и Д. Миддльтон, «Теоретическое сравнение визуального, слухового и инструментального приема импульсных сигналов при наличии шумов», *Journ. Appl. Physics*, 17, 940, ноябрь 1946.

го сигнала можно было бы еще использовать, то мы тем самым пришли бы к абсурдному заключению, поскольку уже отмечалось, что наиболее вероятное значение τ получается при выборе выходного сигнала наибольшей амплитуды независимо от его формы.

5. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Как было показано, распределение апостериорных вероятностей для временного запаздывания τ принимает форму, аналогичную форме модулированных высокочастотных колебаний на выходе линейного фильтра, искаженных по амплитуде или «выпрямленных» устройством с показательной характеристикой. Это, естественно, приводит к такой функции от τ , огибающая которой при достаточно большом сигнале имеет пик вблизи истинного значения τ_0 . Под огибающей этой функции лежит множество тонких пиков, создаваемых несущей частотой.

Эта тонкая структура представляет последовательность вероятных и невероятных значений τ , возникающих при сравнении фазы несущей частоты в функциях $Y(t)$ и $G(t)$. Когда создаваемое этой тонкой структурой весьма ненадежное знание дальности не представляет интереса, она может быть устранена путем сглаживания или «детектирования» апостериорного распределения таким образом, чтобы сохранить площади в интервалах высокочастотных периодов. Действительно, когда τ изменяется со временем достаточно быстро, чтобы воспроизвести данные, заключенные в высокочастотной структуре распределения, от одного периода развертки до следующего, но недостаточно быстро, чтобы это значительно сказалось на модуляции, данные последующих периодов могут быть объединены только после устранения высокой частоты.

Апостериорные распределения после «детектирования» должны быть перемножены, либо произведено сложение их логарифмов. Таким путем однозначно определяется идеальная характеристика детектирования (имеющая форму $\log I_0$, где I_0 — модифицированная функция Бесселя), если после детектирования должно быть произведено последовательное суммирование импульсов.

В настоящей работе не делается попытка развить теорию дальше или устранить из нее принятые допущения. Из сказанного должно быть ясно, что задача получения всех данных из сигнала, искаженного шумами, в принципе может быть однозначно решена одним универсальным методом. Для этого необходимо сформулировать вопрос, написать распределение апостериорных вероятностей для всех возможных ответов на этот вопрос и интерпретировать полученное выражение конкретным физическим устройством. Таким образом, задача расшифровки сигналов исключается, поскольку апостериорное распределение и представляет искомыми данные.

КРАТКИЙ УКАЗАТЕЛЬ ЛИТЕРАТУРЫ

- А. Н. Колмогоров, Основные понятия теории вероятностей, ОНТИ, 1936.
- А. Я. Хинчин, Основные законы теории вероятностей, ГТТИ, 1932.
- Б. В. Гнеденко, Курс теории вероятностей, ГТТИ, 1950.
- Г. Крамер, Случайные величины и распределения вероятностей, ГИИЛ, 1947.
- М. А. Леонтович, Статистическая физика, ГТТИ, 1944.
- А. А. Харкевич, Спектры и анализ, ГТТИ, 1952.
- И. С. Гоноровский, Частотная модуляция и ее применение, Связьиздат, 1948.
- В. Л. Грановский, Электрические флуктуации, ОНТИ, 1936.
- В. И. Бунимович, Флюктуационные процессы в радиоприемных устройствах, Советское радио, 1951.
- А. Я. Хинчин, Теория корреляции стационарных стохастических процессов, „Успехи математических наук“, вып. 5, 1938.
- А. Н. Колмогоров, Статистическая теория колебаний с непрерывным спектром, Юбилейный сборник АН СССР, т. I, 1947.
- В. А. Котельников, О пропускной способности „эфира“ и проволоки в электросвязи, „Материалы к I Всесоюзному съезду по вопросам технической реконструкции связи“, 1933.
- В. А. Котельников, Проблемы помехоустойчивой радиосвязи, „Радиотехнический сборник“, Госэнергоиздат, 1947.
- В. И. Сифоров, О влиянии помех на прием импульсных сигналов, „Радиотехника“, № 1, 1946.
- С. В. Бородич, О помехоустойчивости связи с импульсно-кодовой модуляцией, „Радиотехника“, № 5, 1949.
- В. И. Бунимович и М. А. Леонтович, О распределении числа больших отклонений при электрических флуктуациях, „Доклады АН“, новая серия, вып. 1, 1946.

ОГЛАВЛЕНИЕ

Предисловие	3
Часть I	
К. Шэннон. Статистическая теория передачи электрических сигналов	7
Введение	7
Глава I. Дискретные системы без шумов	11
1. Дискретный канал без шумов	11
2. Источник дискретных сообщений	14
3. Графическое представление цепей Маркова	16
4. Эргодические и смешанные источники	17
5. Выбор, неопределенность и „энтропия“	19
6. „Энтропия“ источника сообщений	23
7. Представление процессов кодирования и декодирования	27
8. Основная теорема для канала без шумов	28
9. Обсуждение результатов и примеры	31
Глава II. Дискретный канал с шумами	35
10. Представление дискретного канала с шумами	35
11. Ненадежность и пропускная способность канала	36
12. Основная теорема для дискретного канала при наличии шумов	39
13. Обсуждение результатов	43
14. Пример дискретного канала	44
15. Пропускная способность канала в некоторых специальных случаях	46
16. Пример эффективного кодирования	48
Глава III. Непрерывные сообщения	49
17. Множества и ансамбли функций	49
18. Ансамбли функций с ограниченной полосой частот	53
19. «Энтропия» непрерывного распределения	54
20. «Энтропия» ансамбля функций	57
21. Потеря «энтропии» в линейных фильтрах	59
22. «Энтропия» суммы двух ансамблей функций	61
Глава IV. Канал с непрерывной передачей	63
23. Пропускная способность канала с непрерывной передачей	63
24. Пропускная способность канала при ограничении средней мощности	65
25. Пропускная способность канала при ограничении пиковой мощности	69
Глава V. Скорость создания сообщений для непрерывного источника	73
26. Функции оценки верности воспроизведения	73
27. Скорость источника при данной оценке верности	76
28. Вычисление скорости создания сообщений	78

Приложение 1	80
Приложение 2	80
Приложение 3	82
Приложение 4	83
Приложение 5	84
Приложение 6	85

Часть II

С. Райс. Теория флуктуационных шумов	88
<i>Глава I.</i> Дробовой эффект	88
1.1. Вероятность поступления на анод точно K электронов за промежуток времени T	88
1.2. Теорема о наложении случайных возмущений	89
1.3. Доказательство теоремы о наложении случайных возмущений	91
1.4. Распределение тока $I(t)$	93
1.5. Обобщение теоремы о наложении случайных возмущений	95
1.6. Сходимость распределения тока I к нормальному закону	99
1.7. Составляющие Фурье тока $I(t)$	100
<i>Глава II.</i> Энергетический спектр и функция корреляции	104
2.1. Некоторые результаты обобщенного гармонического анализа	105
2.2. Энергетический спектр для постоянной и периодической составляющих	106
2.3. Обсуждение выводов первого раздела — ряды Фурье	108
2.4. Обсуждение выводов первого раздела — теорема Парсевала	110
2.5. Гармонический анализ случайных функций	112
2.6. Первый пример — дробовой эффект	114
2.7. Второй пример — случайный телеграфный сигнал	118
2.8. Представление тока шумов	121
2.9. Нормальное распределение нескольких переменных	123
2.10. Центральная предельная теорема	124
<i>Глава III.</i> Статистические свойства флуктуационных шумов	126
3.1. Распределение тока шумов	126
3.2. Распределение $I(t)$ и $I(t+\tau)$	129
3.3. Ожидаемое число нулей в 1 сек.	131
3.4. Распределение нулей	137
3.5. Кратные интегралы, входящие в формулы	147
3.6. Распределение выбросов тока шумов	151
3.7. Плотность вероятностей для огибающей тока шумов	155
3.8. Выбросы огибающей	159
3.9. Флуктуации энергии	167
3.10. Распределение суммы из тока шумов и синусоидального тока	178
3.11. Некоторые дополнения о токах дробового эффекта	187
<i>Глава IV.</i> Прохождение сигнала и шумов через нелинейные устройства	189
4.1. Ток низкой частоты на выходе квадратичного детектора	189
4.2. Ток низкой частоты на выходе линейного детектора	192
4.3. Некоторые статистические свойства тока на выходе нелинейного устройства общего типа	196
4.4. Выходной энергетический спектр	204

4.5. Энергетический спектр на выходе устройства с квадратичной характеристикой	205
4.6. Два метода функции корреляции	211
4.7. Линейное детектирование шумов — первый метод	212
4.8. Метод характеристической функции	215
4.9. Энергетический спектр на выходе нелинейного устройства общего типа	218
4.10. Некоторые результаты, полученные методом функции корреляции	223
Приложение 4 А. Таблица нелинейных устройств, описываемых интегралами	230
Приложение 4 В. Функция ${}_1F_1(a; c; x)$	231
Приложение 4С. Энергетический спектр, соответствующий ψ^2	233

Часть III

Применение статистических методов в радиолокации	239
<i>Глава I.</i> Ф. Вудворд и И. Дэвис. Принцип «обратной вероятности» в теории передачи сигналов	239
1. «Обратная вероятность» и передача сообщений	239
2. Теория получения данных	243
3. Шумы и апостериорное распределение	248
4. Корреляционный приемник	251
5. Заключение и обсуждение результатов	254
Приложение	256
<i>Глава II.</i> Ф. Вудворд и И. Дэвис. Статистическая теория приема радиолокационных сигналов	258
1. Введение	258
2. Математическое представление колебаний	260
3. Основы теории радиолокационного наблюдения	262
4. Функции сигнала и шумов	266
5. Точность измерения дальности	268
6. Надежность наблюдения	271
7. Количество данных при радиолокационном наблюдении	273
8. Обсуждение результатов	275
<i>Глава III.</i> Ф. Вудворд. Проектирование радиолокационных приемников на основе статистических методов	278
1. Введение	278
2. Принцип «обратной вероятности» и функция количества данных	279
3. Вычисление функций количества данных	280
4. Идеальный фильтр, включенный перед детектором	282
5. Обсуждение результатов	284
<i>Краткий указатель литературы</i>	285

Редактор Л. Б. БАСКАКОВА

Технический редактор А. В. Виленева Корректор А. С. Кириллова

Сдано в производство 22/XII 1952 г. Подписано к печати 17/III 1953 г. А02085.

Бумага 60×92¹/₁₆—9 бум. л.—18 печ. л. Уч.-изд. л. 18,5. Изд. № 15/1981.

Цена 14 р. 45 к. Зак. № 974.