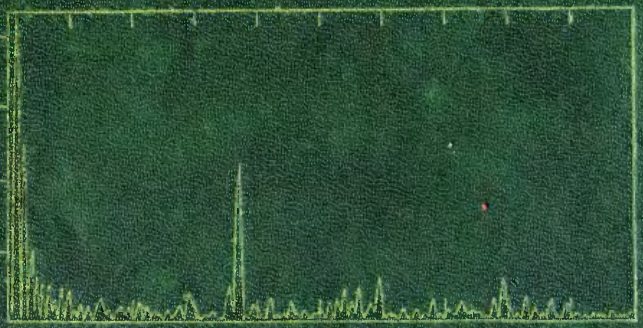
Б19Р
Т35

В. Ю. ТЕРЕБИЖ

АНАЛИЗ ВРЕМЕННЫХ РЯДОВ В АСТРОФИЗИКЕ



19
Т 2 1

В. Ю. ТЕРЕБИЖ

АНАЛИЗ ВРЕМЕННЫХ РЯДОВ В АСТРОФИЗИКЕ

573412

МОСКВА «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
1992

ББК 22.63
Т35

УДК 524.314

Теребиж В. Ю. Анализ временных рядов в астрофизике.-М.: Наука. Гл. ред. физ.-мат. лит., 1992. - 392 с. -
ISBN 5-02-014484-3.

Рассматриваются методы анализа последовательностей измерений и, в частности, выявление скрытой периодичности и счет фотонов. Наряду с классическим подходом к спектральной оценке дается метод максимальной энтропии и новые результаты теории неравномерных временных рядов. Общие положения иллюстрируются многочисленными примерами - как модельными, так и взятыми из астрономической практики. Приводятся необходимые расчетные формулы и алгоритмы. Значительная часть материала впервые дается в монографическом изложении.

Для астрономов, физиков и математиков - специалистов, аспирантов и студентов. Может быть полезна исследователям других областей (биологам, инженерам), связанным с изучением внутренней структуры временных процессов.

Табл. 9. Ил. 68. Библиогр. 348 назв.

1605040000 - 017
Т——————— 104-92
053(02)-92

© "Наука". Физматлит, 1992

ISBN 5 - 02 - 014484 - 3

О Г Л А В Л Е Н И Е

Предисловие.....5
Глава 1. **Основные понятия**.....9
§ 1.1. Вводный пример.....11
§ 1.2. Постановка задачи.....25
§ 1.3. Интеграл Фурье.....29
§ 1.4. Периодические функции.....37
§ 1.5. Спектральный анализ стационарных случайных функций.....45
§ 1.6. Реальные процессы.....56
§ 1.7. Два подхода к проблеме поиска периодичности.....69
Глава 2. **Равномерные временные ряды**.....75
§ 2.1. Основные соотношения.....76
§ 2.2. Теорема отсчетов.....84
§ 2.3. Статистические свойства периодограммы.....87
§ 2.4. Сглаживание периодограммы.....100
§ 2.5. Спектральное оценивание с помощью метода максимума энтропии.....109
§ 2.6. Модели авторегрессии.....125
§ 2.7. Вариации солнечной активности.....141
§ 2.8. Анализ кривой блеска квазара 3С 273.....146
§ 2.9. Взаииная спектральная плотность двух процессов.....160
Глава 3. **Спектральная теория неравномерных временных рядов**.....167
§ 3.1. LS_m- спектр.....167
§ 3.2. Случайная последовательность моментов наблюдения.....177
§ 3.3. Отбеливание временных рядов.....184
§ 3.4. Алгоритм **CLEAN**.....201
§ 3.5. Выделение тренда.....204
§ 3.6. Фликкер-шум. Переменность блеска ядер сейфертовских галактик.....212
§ 3.7. Аналогии в оптике.....227
§ 3.8. Некоторые общие методы решения обратных задач.....238

Глава 4. **Непараметрические методы выделения скрытой периодичности**.....250
§ 4.1. Статистика Аббе-Лафлера-Кинмана.....251
§ 4.2. Фазовая группировка данных (методы Уиттекера - Г. Робинсона, Уорнера - Э. Робинсона и Стеллингвефа).....254
§ 4.3. Методы, сохраняющие эффективность при малом количестве измерений («натянутой струны», Ренсона и Пельта).....260
§ 4.4. Автокорреляционные методы.....266
§ 4.5. Сопряженные периоды.....268
§ 4.6. Сопоставление различных методов поиска скрытой периодичности.....273
Глава 5. **Импульсные случайные последовательности**.....278
§ 5.1. Пуассоновский процесс.....279
§ 5.2. Спектральный анализ импульсных последовательностей.....293
§ 5.3. Методы, опирающиеся на группировку событий...306
§ 5.4. Примеры.....314
§ 5.5. Счет фотонов.....324
§ 5.6. Эффекты мертвого времени. Отношение сигнала к шуму при счете событий.....335
Заключение.....348
Приложения.....350
Приложение 1. Значения некоторых сумм.....350
Приложение 2. Основные формулы с δ-функциями Дирака.....350
Приложение 3. Свойства функций $\mathrm{sinc}(x)$, $\phi_N(x)$ и $\mathrm{tr}(x)$...352
Приложение 4. Характеристики случайных величин.....353
Приложение 5. Процентные точки для отношения максимального значения периодограммы к среднему.....359
Приложение 6. Процентные точки для критерия Аббе.....360
Приложение 7. Числовые данные.....361
Список литературы.....365
Указатель иностранных авторов.....382
Предметный указатель.....389

Посвящаю памяти М. А. Аракеляна

ПРЕДИСЛОВИЕ

Понятие *временного ряда* как *упорядоченной совокупности дискретных измерений* включает обширный круг данных наблюдений в физике, экономике и технике. К временным рядам можно отнести, например, последовательно измеренные значения координаты броуновской частицы или интенсивности звучания музыкального инструмента, данные о производстве какого-либо продукта в разные годы, индексы солнечной активности, записанный на микрофотометре спектр источника излучения, оценки блеска переменной звезды, совокупность моментов регистрации фотонов и т.д. Характерным свойством реальных временных рядов является наличие случайного компонента, иногда представляющего собой просто шум, но нередко выступающего и в качестве основного источника информации о физическом процессе.

Два фактора обусловили широкий интерес к временным рядам, проявляющийся в последние годы в астрофизике. Во-первых, по мере повышения точности измерений выяснилось, что такие характеристики большинства астрономических объектов как блеск, лучевая скорость быстро изменяются со временем, и традиционных методов недостаточно для анализа таких вариаций. Во-вторых, все большее распространение получает счет отдельных событий - фотонов в оптической, рентгеновской и γ-астрономии, импульсов в детекторах гравитационных волн и нейтрино - и вообще изучение процессов, где форма и амплитуда отклика не важны, а физическая информация содер-

жится во временном распределении отсчетов. Следует отметить также, что бурное развитие стохастической динамики показало фундаментальное значение вероятностного описания не только в квантовой, но и в классической физике.

Целью анализа временных рядов является выяснение механизма упорядочения отсчетов, что обычно выражается в построении формальной модели изучаемого процесса. Так, в простейших случаях интерес может представить определение относительно медленного изменения членов ряда - тренда или поиск периодического компонента, неявно присутствующего в рассматриваемом явлении. Более общая постановка задачи предполагает восстановление из данных наблюдений *спектра* временных колебаний процесса и последующее уточнение его внутренней структуры.

Временным рядам уделяется много внимания, и тем не менее явно ощущается недостаток сведений, необходимых для практической работы. Причина заключается в том, что имеющиеся монографии излагают почти исключительно строгую теорию равномерных временных последовательностей, которая, конечно, не может охватить многих прикладных задач. В частности, для астрономических наблюдений характерны неравномерность распределения моментов наблюдения во времени, привнесение ложных периодичностей, обусловленных суточным, лунным и годовым циклами, и важная роль эффектов, связанных с квантовой природой излучения.

Уже первая фундаментальная работа по теории временных рядов - работа Лагранжа (1873) - была выполнена в связи с потребностями астрономии. Эти же требования побудили ввести многие важные понятия теории временных рядов, например, периодограмму (А.Шустер, 1898, 1906) и авторегрессионную модель (Г.Юл, 1927). Между тем, сложилось так, что результаты современной теории, которыми широко пользуются радиоинженеры, геофизики и биологи, остаются малоизвестными астрономам, ограничивающимся зачастую лишь определением периода колебаний блеска или лучевой скорости объекта.

В данной книге помимо традиционных методов спектрального оценивания достаточно подробно рассматриваются методы анализа импульсных и неравномерных временных рядов с учетом

перечисленных выше особенностей астрономических измерений (которые, впрочем, в различной мере присуши и другим дисциплинам). Следует подчеркнуть, что проблема заключается не столько в разработке методической стороны учета группировки данных во времени, сколько в выяснении способов извлечения информации из наблюдений стохастического *по своей природе* процесса, например, последовательности моментов регистрации отдельных фотонов или изменений светимости объекта, аккрецирующего газ.

Мы не стремились максимально полно осветить имеющуюся литературу. Вместе с тем важнейшие понятия и методы по возможности рассматриваются с различных точек зрения. Значительная часть материала впервые приводится в монографическом изложении, некоторые результаты являются новыми. Астрономическая тематика большинства примеров, разумеется, не ограничивает общности методики.

Как заметил известный специалист в области анализа временных рядов Джейнс (1982), «... в настоящее время написать программу для компьютера и выполнить на нем соответствующие вычисления гораздо проще, чем серьезно обдумать суть проблемы». Поскольку имеются руководства Энслейна и др. (1977), Пельта (1980), Пресса и др. (1986), в которых подробно изложены численные методы анализа временных рядов, мы предпочли ввести как можно больше примеров, для каждого из которых даны исходные значения ряда, расчетные соотношения и результаты анализа в виде графиков. Это позволяет при необходимости составить и проверить программу вычислений.

Предполагается, что читателю известны основы теории вероятностей (краткая сводка используемых соотношений дана в Приложении 4). Сведения из теории случайных процессов приводятся по мере надобности, однако было бы весьма желательным знакомство с превосходными курсами Рытова (1976) или Дженкинса и Ваттса (1968).

Принятая нумерация формул указывает последовательно номера главы, параграфа и формулы в соответствующем параграфе; при ссылках в пределах одного параграфа указывается только последний номер.

Значительное влияние на характер изложения в данной кни-

ге оказали, помимо отмеченных выше, работы Брейсуэлла (1965), Диминга (1975), Фелжетта (1972), Фейнмана и др. (1963), Кокса и Льюиса (1966), Крамера (1946), Ланцоша (1956), Пресса (1978) и Робинсона (1982).

Большинство численных расчетов выполнила А.В.Теребиж. Много полезных замечаний высказали В.В.Бирюков, В.П.Маланушенко, Я.Пельт и С.М.Рытов, ознакомившиеся с отдельными частями книги. Всем им автор приносит искреннюю благодарность.

ГЛАВА 1

ОСНОВНЫЕ ПОНЯТИЯ

Классический способ описания физических явлений основан на представлении изучаемых характеристик непрерывными функциями времени или, как часто предпочитают говорить, непрерывными временны́ми процессами. В случае, когда процесс нестационарен, сначала исследуется его среднее, систематическое изменение во времени. Однако нестационарность может отражать лишь недостаточную стабильность условий эксперимента и не представлять интереса сама по себе, либо быть связанной с физическими явлениями, отличными от исследуемого. Так мы приходим к понятию *стационарного временно́го процесса*, применяющемуся для описания состояний, средние свойства которых сохраняются достаточно долго.

Временно́е описание процессов, будучи полным, не является, вообще говоря, ни единственным, ни наиболее удобным. Существует иное - спектральное описание, определяющее характер *типичных* временных изменений процесса. Особенно полезен спектральный подход при изучении стационарных процессов, когда конкретные временные реализации процесса далеко не очевидным образом характеризуют его внутреннюю структуру. Так, знанию временных флуктуаций напряженности электромагнитного поля в какой-либо точке пространства физик предпочтет сведения о распределении энергии колебаний по частотам - спектральной плотности или просто спектре процесса. Особенно привычным является спектральное представление излучения в оптическом диапазоне, т.е. света. Столь же информативным это представление оказывается и при изучении вариаций с произвольным характерным масштабом времени, и в

частности, изменений потока излучения от нестационарных астрономических объектов; как и в привычном нам спектре излучения газа, гармоническим составляющим отвечают узкие спектральные линии, а хаотическим колебаниям - непрерывный спектр. Очевидно, что для построения физической модели процесса изучение непрерывного спектра необходимо наряду с выделением спектральных линий (скрытой периодичности).

В этой книге рассматривается, главным образом, направление, связанное с восстановлением полного спектра временных вариаций, а также построением на этой основе соответствующей модели процесса.

Всякий реальный прибор имеет конечную постоянную времени, поэтому невозможно точно воспроизвести параметры непрерывного временного процесса: экспериментальные отсчеты по необходимости дискретны. Необратимость процессов во времени естественным образом упорядочивает всякий процесс измерений, и в тех случаях, когда порядок имеет значение, мы приходим собственно к временному ряду. Однако упорядоченность может оказаться существенной не только при исследовании развития процесса во времени; тогда в качестве t выступает пространственная координата или номер измерения. Достаточно современным примером может служить процесс регистрации изображений с помощью дискретного набора приемников излучения. Проблемы, возникающие при интерпретации таких рядов, не зависят от конкретного смысла параметра t. Для определенности t почти везде далее рассматривается как временная координата.

Обыкновенно временные ряды порождаются дискретностью отсчетов какого-либо непрерывного процесса, но возможны и ситуации, когда дискретность вызвана самой постановкой задачи. Предположим, например, что изучается количество энергии, выделяющейся при вспышках на поверхности звезды. Вспышки случаются достаточно редко; если мы хотим обнаружить систематическое изменение активности звезды, следует суммировать энергии отдельных вспышек за определенные промежутки времени и проанализировать получающийся таким образом дискретный ряд.

Особый раздел теории временных рядов составляет изучение

процессов, дискретность которых обусловлена квантовой природой вещества и, следовательно, в большей мере отражает не возможности наблюдателя, а суть исследуемых явлений. Таким рядам посвящена гл. 5.

Ограничимся пока этими необходимыми пояснениями и рассмотрим в § 1.1 простой пример временного ряда, иллюстрирующий некоторые характерные особенности предмета. Это позволит нам в дальнейшем более конкретно обсудить постановку задачи анализа временных рядов (§ 1.2) и очертить круг вспомогательных понятий (§ 1.3-1.6).

§ 1.1. Вводный пример

Данные для анализа. На рис. 1.1.1 и в табл. П.7.1 (см. Приложение 7) представлен ряд $\{y_k\}$, называемый в дальнейшем для краткости рядом A. Он включает $N = 50$ отсчетов, которые соответствуют моментам времени

$$t_k = k\Delta t, \quad k = 0, 1, 2, \ldots, N - 1. \tag{1.1.1}$$

Ряд A относится к классу *равномерных* временных рядов, для которых, в отличие от неравномерных рядов, интервалы времени между последовательными отсчетами равны одному и тому же значению Δt. Следующие три вопроса возникают уже при первом знакомстве с каждым временным рядом: 1) содержит ли ряд систематическое изменение - тренд $f_0(t_k)$? 2) не является ли освобожденный от тренда ряд $\{y(t_k)-f_0(t_k)\}$ набором независимых друг от друга случайных чисел $\xi(t_k)$? 3) можно ли получить объективные указания на существование периодического компонента $f_p(t_k)$? Рассмотрение рис. 1.1.1 показывает, что наличие незначительного тренда для ряда A исключить нельзя. Далее, кажется маловероятным, что вторая из перечисленных выше возможностей реализуется в чистом виде, т.е. ряд является случайным шумом, однако наличие шума в качестве аддитивного компонента вполне возможно. Наконец, достаточно правдоподобной представляется и реальность периодического компонента.

Таким образом, мы допускаем, что процесс, порождающий ряд A, включает следующие три компонента:

$$y(t) = f_0(t) + f_p(t) + \xi(t). \qquad (1.1.2)$$

Следует сразу оговорить, что представление конкретного ряда в виде (2) неоднозначно (столь же неоднозначно, впрочем, и любое другое представление). Непросто также дать и строгие определения для каждого из входящих в (2) слагаемых. В частности, насколько быстрые изменения допустимы для тренда? Как отличить реальные тренд и периодичность от аналогичных случайных явлений, обязательно присутствующих в конечной реализации шума? Вопросы подобного рода можно продолжить; чтобы избежать излишнего усложнения задачи на вводной стадии, мы ограничимся пока интуитивными соображениями.

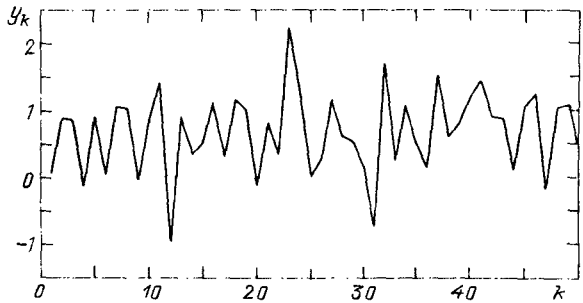


Рис. 1.1.1. Значения $\{y_k\}$ для временного ряда A

Тренд. Для оценки характера тренда целесообразно аппроксимировать наблюдаемые значения $\{y(t_k)\}$ полиномами невысокой степени от t_k, начиная с вычисления среднего

$$\bar{y} = \frac{1}{N} \sum_{k=0}^{n-1} y_k, \qquad (1.1.3)$$

линейной зависимости

$$y_k = a_0 + a_1 t_k, \quad k = 0, 1, \ldots, N - 1 \qquad (1.1.4)$$

и т.д. Аппроксимация методом наименьших квадратов дает для ряда A:

$$\bar{y} = 0,321 \pm 0,086,$$

$$a_0 = 0,177 \pm 0,169, \qquad (1.1.5)$$

$$a_1 = 0,0059 \pm 0,0059,$$

что согласуется с предварительным выводом о незначительности тренда. Дальнейшее изучение тренда может включать скользящее усреднение отсчетов, либо усреднение в неперекрывающихся интервалах времени. При этом следует ожидать, что вклад случайного и периодического компонентов уменьшится, в то время как систематического - станет более контрастным. Результаты такого усреднения для ряда A не усиливают указаний на реальность тренда с меньшим характерным масштабом времени, и мы примем в качестве первого приближения, что тренд в данном случае отсутствует.

Периодограмма. Каким образом можно было бы выделить периодический компонент $f_p(t)$? Представляются естественными следующие соображения. Практически всякую периодическую функцию $f_p(t)$ с периодом P_0 можно разложить в ряд Фурье

$$f_p(t) = A_0 + \sum_{n=1}^{\infty} A_n \cos(2\pi n \nu_0 t - \varphi_n), \qquad (1.1.6)$$

т.е. представить в виде суммы простых гармонических колебаний с равноотстоящими частотами $\nu_n = n\nu_0$, кратными основной частоте $\nu_0 = P_0^{-1}$. Поэтому наблюдаемый временной ряд $\{y_k\}$ можно сравнить с пробными гармоническими колебаниями

$$\tilde{f}(t) = \cos(2\pi\nu t - \varphi), \qquad (1.1.7)$$

и в тех случаях, когда частота ν и фаза φ пробного колебания будут близки соответствующим значениям (ν_n, φ_n) одной из гармоник $f_p(t)$, следует ожидать, что определенная надлежащим образом степень согласия между рядами $\{y_k\}$ и $\{\tilde{f}(t_k)\}$ будет максимальна. Ситуация здесь аналогична исследованию сложного аккорда при помощи набора пробных камертонов.

В качестве меры близости рядов $\{y_k\}$ и $\{\tilde{f}(t_k)\}$ примем абсолютное значение (или квадрат) величины

$$C_N(\nu,\varphi) = \frac{1}{N}\sum_{k=0}^{N-1}\cos(2\pi\nu t_k - \varphi)y_k^{(0)}, \qquad (1.1.8)$$

где $\{y_k^{(0)}\}$ – центрированный ряд:

$$y_k^{(0)} = y_k - \overline{y}, \qquad \sum_{k=0}^{N-1} y_k^{(0)} = 0, \qquad (1.1.9)$$

Если пробная синусоида $\tilde{f}(t)$ такова, что значения $\tilde{f}(t_k)$ положительны преимущественно тогда же, когда положительны и значения $y_k^{(0)}$, а отрицательные значения $\tilde{f}(t_k)$ чаще всего сочетаются с отрицательными же значениями $y_k^{(0)}$, то коэффициент C_N будет (по модулю) велик. Описанная ситуация означает, что между $y(t_k)$ и $\tilde{f}(t_k)$ имеет место хорошее согласие. При беспорядочном сочетании колебаний $\tilde{f}(t)$ и $y(t)$ ожидаемое значение C_N близко к нулю. Параметр C_N, отличающийся от коэффициента корреляции лишь постоянным множителем, в той же степени, что и последний, может служить мерой линейной зависимости последовательностей $\{y_k\}$ и $\{\tilde{f}_k\}$.

В соответствии со сказанным выше следует найти те значения ν и φ, при которых $C_N^2(\nu,\varphi)$ достигает максимума. Варьируем сначала φ и определим значение $\varphi_0(\nu)$, доставляющее максимум C_N^2 при заданной частоте ν. Иными словами, станем «двигать» пробную синусоиду заданной частоты ν вдоль оси времени до тех пор, пока не добьемся наилучшего ее согласия с рядом $\{y_k\}$.

Перепишем (8) в виде

$$C_N(\nu,\varphi) = U_N(\nu)\cos\varphi + V_N(\nu)\sin\varphi, \qquad (1.1.10)$$

где

$$U_N(\nu) = \frac{1}{N}\sum_{k=0}^{N-1}\cos(2\pi\nu k\Delta t)y_k^{(0)},$$

$$V_N(\nu) = \frac{1}{N}\sum_{k=0}^{N-1}\sin(2\pi\nu k\Delta t)y_k^{(0)}. \qquad (1.1.11)$$

Из условия $\partial C_N/\partial\varphi = 0$ нетрудно найти, что

$$\operatorname{tg}\varphi_0(\nu) = V_N(\nu)/U_N(\nu), \qquad (1.1.12)$$

а подстановка $\varphi_0(\nu)$ в (10) дает следующее выражение для максимального при данной частоте квадрата коэффициента C_N:

$$C_N^2\ [\nu,\ \varphi_0(\nu)] \equiv P_N(\nu) = U_N^2(\nu) + V_N^2(\nu). \qquad (1.1.13)$$

Определить аналитически те значения частоты, при которых $P_N(\nu)$ достигает максимума, значительно сложнее, чем подобрать подходящую фазу. На практике можно просто вычислить все значения $P_N(\nu)$ в интересующем нас интервале частот (периодов) с некоторым шагом $\delta\nu$, и наиболее выделяющиеся максимумы укажут на частоты вероятных гармоник, содержащихся во временном ряду $\{y_k\}$.

Функция $P_N(\nu)$, обычно называемая *периодограммой*, была введена с целью выявления скрытых периодичностей А. Шустером (1898, 1906). Изучение периодограммы Шустера до сих пор является одним из наиболее распространенных средств при анализе равномерных и близких к таковым временных рядов. Первоначально в качестве аргумента P_N рассматривался пробный период ν^{-1}, что и обусловило название этой функции. В настоящее время термин «периодограмма» представляется не вполне удачным, поскольку анализ обычно проводится в частотной области, где достаточна равномерная сетка пробных частот (см. § 1.6).

Периодограмма Шустера для ряда А представлена на рис. 1.1.2. Как это обычно делается при рассмотрении равномерных рядов, единицей измерения времени выбран промежуток Δt, а единицей измерения частоты – соответственно Δt^{-1}. Значения $P_N(\nu)$ вычислены с равномерным шагом по частоте $\delta\nu = 1/500$. Нетрудно заметить, что график $P_N(\nu)$ в промежутке $0,5 \leqslant \nu \leqslant 1,0$ представляет собой отражение графика в области $0 \leqslant \nu \leqslant 0,5$ относительно прямой $\nu = 0,5$, а область $0 \leqslant \nu \leqslant 1,0$ повторяется на всех промежутках $n \leqslant \nu \leqslant n+1$, $n = 1, 2, \ldots$ Таким образом, если принять во внимание симметрию периодограммы

$$P_N(-\nu) = P_N(\nu), \qquad (1.1.14)$$

то можно заключить, что информация о частотной структуре ряда содержится в промежутке частот $0 \leqslant \nu \leqslant 0,5$. В размерных единицах граничная частота равна

$$\nu_c = \frac{1}{2\Delta t}, \qquad (1.1.15)$$

она называется *частотой Найквиста* (Найквист, 1924, 1928). Понятие критической частоты играет фундаментальную роль в теории равномерных временных рядов и с определенными оговорками может применяться и при рассмотрении рядов, близких к равномерным. Описанная симметрия периодограммы относительно ν_c и ее повторяемость на целых интервалах в общем случае нарушаются тем сильнее, чем значительнее отличается ряд от равномерного. Условие же (14) справедливо для произвольных действительных рядов.

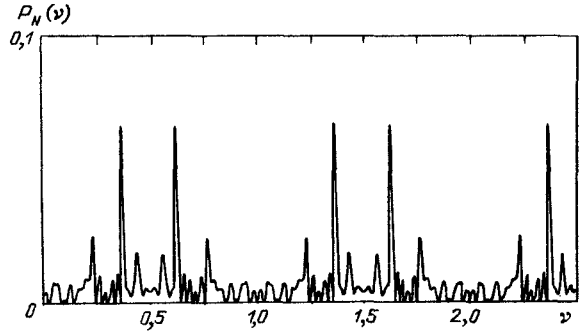


Рис. 1.1.2. Периодограмма $\{y_k\}$ ряда А

Подмена частот. В соответствии со смыслом периодограммы максимумы $P_N(\nu)$ отвечают вероятным гармоникам, содержащимся в исследуемом процессе. Однако, как свидетельствует рис. 1.1.2, лишь часть максимумов соответствует реальным гармоническим колебаниям, остальные максимумы являются, можно сказать, «духами» истинных. Вместе с тем несомненно, что гармоники с ложными частотами столь же удовлетворительно согласуются с исследуемым рядом, как и реальные колебания, поскольку максимумы $P_N(\nu)$ для тех и других имеют одинаковую интенсивность. Описанное явление называется *подменой частот* (aliasing), а отвечающие ложным максимумам частоты – *сопряженными*, *неразличимыми*, *ложными* или *частотами подмены*. Мы будем пользоваться преимущественно первым из этих терминов. В исследованиях конкретных рядов с невысоким

отношением сигнала к шуму и моментами наблюдений, модулированными некоторой частотой (т.е. при условиях, типичных для астрономических наблюдений с их суточной, лунной, сезонной и годовой модуляциями), выделение истинных периодов составляет одну из труднейших частей анализа. На протяжении этой книги мы неоднократно будем возвращаться к явлению подмены частот, рассматривая его в различных аспектах (например, достаточно наглядной представляется интерпретация, связывающая это явление с образованием спектров различных порядков дифракционной решеткой). Пока же отметим следующее.

Предположим, что истинная периодическая зависимость $\{f_p(t_k)\}$, входящая в исследуемый ряд $\{y_k\}$, имеет период P_0 и основную частоту $\nu_0 = P_0^{-1}$. Тогда с $\{y_k\}$ будет хорошо согласовываться пробное колебание вида (7) с частотой ν_0, т.е. $\cos(2\pi\nu_0 t_k - \varphi)$, и на частоте ν_0 периодограмма имеет (истинный) максимум. Однако ввиду дискретности моментов наблюдения $t_k = k\Delta t$ справедливо равенство

$$\cos(2\pi\nu_0 t_k - \varphi) = \cos\left[2\pi(\nu_0 \pm \frac{l}{\Delta t})t_k - \varphi\right], \qquad (1.1.16)$$

где $l = 0, 1, 2, \ldots$ - произвольное целое число. Это означает, что дискретный ряд наблюдений с шагом Δt аппроксимируется не только гармоникой с истинной частотой ν_0, но и любой из гармоник с частотами $\nu_0 \pm l/\Delta t$. Явление подмены частот хорошо знакомо по кинофильмам, где вращение, скажем, колес повозки кажется происходящим с частотой, отличной от действительной. Иногда можно видеть, что колеса вращаются в противоположную сторону (частота подмены отрицательна). Обусловлено это явление, разумеется, дискретностью моментов наблюдения периодического процесса.

Поскольку периодограмма является симметричной функцией частоты, совокупность частот, сопряженных данной частоте ν_0, можно записать в виде

$$\nu_l = \pm\nu_0 \pm \frac{l}{\Delta t}, \quad l = 0, 1, 2, \ldots \qquad (1.1.17)$$

Фазовая диаграмма. Вернемся к рассмотрению ряда А. В области $\nu \leqslant \nu_c$ периодограмма содержит основной максимум на частоте $\hat{\nu}_0 \simeq 0,3705$ и меньшие пики на частотах $\nu' \simeq 0,228$,

$\nu'' \simeq 0,437$ и т.д. Предположим, что основной максимум отвечает реальному периоду $\hat{P}_0 = (0,3705)^{-1} = 2,6991$. Форму этого периодического колебания можно определить на основании следующих простых рассуждений.

Если исследуемая функция $y(t)$ есть сумма некоторой периодической зависимости $f_p(t)$ с периодом P_0 и шума $\xi(t)$, то, накладывая отрезки ряда $\{y_k\}$ длиной P_0 друг на друга, можно выявить $f_p(t)$ наиболее контрастным образом на фоне шума. Поэтому свернем временной ряд с периодом $\hat{P}_0$ и рассмотрим так называемую *фазовую диаграмму* для этого периода. Формально с этой целью следует каждому моменту наблюдения t_k сопоставить нормированную к промежутку [0,1] фазу

$$x_k = \mathrm{FR}\left(\frac{t_k - t_*}{\hat{P}_0}\right), \qquad k = 0, 1, \ldots, N-1, \qquad (1.1.18)$$

где FR(z) означает дробную часть числа z, t_* - произвольно выбранный момент времени, а затем рассмотреть зависимость y_k от x_k. Фазовая диаграмма ряда A для частоты 0,3705 представлена на рис. 1.1.3, где область изменения фазы повторена для наглядности дважды, как это принято у исследователей переменных звезд.

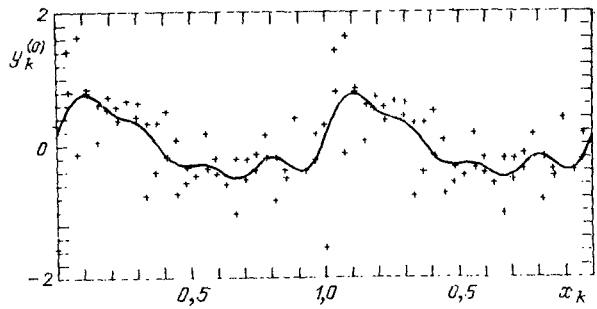


Рис. 1.1.3. Фазовая диаграмма $\{y_k^{(0)}\}$ ряда A для частоты ν=0,3705. Область изменения фазы для наглядности повторена

Мы видим, что искомая периодическая зависимость действительно «проявляется» на фазовой диаграмме. Для лучшего ее выделения проведем при помощи метода наименьших квадратов среднюю «кривую блеска». В качестве аппроксимирующей функ-

ции удобно выбрать тригонометрический полином

$$g_m(x) = \frac{a_0}{2} + \sum_{n=1}^{m} \left[a_n \cos(2\pi n x) + b_n \sin(2\pi n x) \right], \qquad (1.1.19)$$

поскольку фактически фазовая диаграмма задана на цилиндрической поверхности и применение тригонометрических функций автоматически обеспечивает непрерывность полинома $g_m(x)$ и его производных при идентичных по смыслу фазах $x = 0$ и $x = 1$. Процедура, описанная в § 3.3, дает при $m = 4$ значения коэффициентов a_n и b_n, приведенные в третьем и четвертом столбцах табл. 1.1.1.

Таблица 1.1.1

n	ν_n	(a_n)	(b_n)	a_n	b_n	$A_n^2/4$
0	–	0,6164	0,	0,7090	0,	–
1	0,3679	0,2519	0,4469	0,3793	0,2657	0,0536
2	0,7358	-0,0735	0,1638	0,1347	0,0986	0,0070
3	1,1036	-0,0473	0,0883	0,0731	0,0066	0,0013
4	1,4715	-0,0021	0,1244	0,0365	-0,0161	0,0004

Отбеливание. Итак, вполне вероятно, что мы нашли периодический процесс $f_p(t)$, содержащийся в $y(t)$. Представляет интерес вычесть колебание $g_4(x_k)$ из наблюдаемых значений y_k и найти периодограмму ряда остатков $\{y_k - g_4(x_k)\}$. Соответствующий график приведен на рис. 1.1.4. Сравнивая его с рис. 1.1.2, мы видим, что основной максимум $\hat{\nu}_0 = 0,3705$ и все сопряженные ему максимумы (aliases) исчезли. Описанная операция называется *отбеливанием* ряда с частотой $\hat{\nu}_0$.

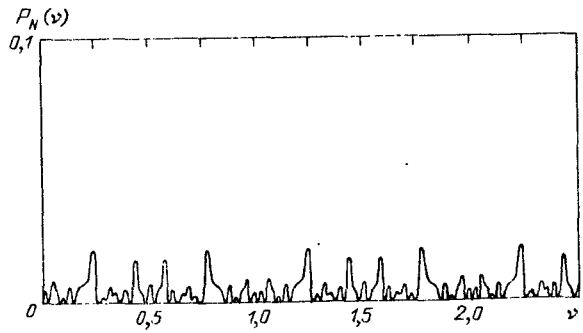


Рис. 1.1.4. Периодограмма остатков $\{y_k - g_4(x_k)\}$ ряда A

Если отвлечься от основного и сопряженного ему максимумов, то общие черты периодограммы сохранились после операции отбеливания. В частности, на рис. 1.1.4 мы вновь видим максимумы с частотами $\nu' \simeq 0,228$, $\nu'' \simeq 0,437$ и др. Отвечают ли они наличию других, помимо найденного, периодических компонентов? Если свернуть ряд остатков $\{y_k - g_4(x_k)\}$ с периодами $(\nu')^{-1}$, $(\nu'')^{-1}$, то расположение точек на фазовых диаграммах мало отличается от хаотического. Однако степень упорядоченности точек на фазовой диаграмме не может служить независимым от периодограммы свидетельством реальности данного периода. По этой причине объективная оценка реальности периодических компонентов должна основываться на уровнях значимости максимумов на периодограмме, т.е. на расчете вероятностей случайного образования максимумов, по крайней мере столь же выделяющихся, как это найдено при рассмотрении конкретных данных. Соответствующая теория будет изложена в § 2.3, а пока нам придется приостановиться в исследовании ряда A.

Суммируем полученные для ряда A результаты. Колебание с периодом $\hat{P}_0 = (0,3705)^{-1} = 2,6991$, по-видимому, единственно; его форма приближенно определяется выражением (19) при значениях коэффициентов, данных в третьем и четвертом столбцах табл. 1.1.1. Достоверность других периодических компонентов, содержащихся в A, значительно уступает достоверности основного компонента.

Сопоставление с исходными данными. Теперь представляется уместным обратиться к достоверной информации о ряде A. Дело в том, что этот ряд имеет искусственное происхождение: он был образован с целью показать эффективность анализа при помощи одного из наиболее распространенных на практике методов. При этом, разумеется, желательно, чтобы внутренняя структура ряда была известна заранее.

Итак, значения ряда A были вычислены при помощи соотношений

$$y(t_k) = f(t_k) + \xi(t_k), \quad t_k = k\Delta t,$$

$$\Delta t = 1, \quad x_k = \mathrm{FR}(t_k/P_0), \quad k = 0, 1, \ldots, 49, \qquad (1.1.20)$$

где периодический компонент

$$f(x) = e^{-\left(\frac{x}{c_1}\right)^2} + e^{-\left(\frac{1-x}{c_2}\right)^2}, \quad 0 \leqslant x \leqslant 1, \qquad (1.1.21)$$

постоянные $c_1 = 0,3$, $c_2 = 0,1$, и значение периода принято равным основанию натуральных логарифмов: $P_0 = e \simeq 2,71828$. Случайный компонент

$$\xi(t_k) = \sigma\, \eta_k, \qquad (1.1.22)$$

где постоянная $\sigma = 1/2$, а η_k суть взаимно независимые нормальные случайные величины с нулевым средним и единичной дисперсией, так что дисперсия шума $\sigma_\xi^2 = \sigma^2 = 1/4$. Использованная реализация $\xi(t_k)$, а также значения исходного $\{f_k\}$ и зашумленного $\{y_k\}$ рядов приведены в табл. П.7.1.(стр. 261)

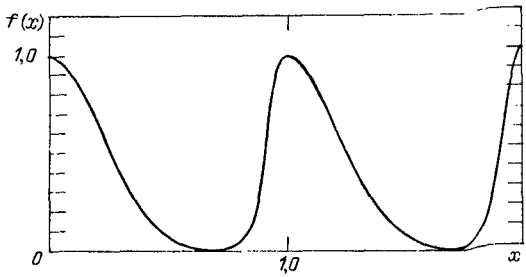


Рис. 1.1.5. Форма периодического компонента $f(x)$ ряда А

Форму исходной периодической зависимости $f(x)$ иллюстрирует рис. 1.1.5. Принятое для $f(x)$ выражение (21) не свободно от небольших дефектов: значения $f(0)$ и $f(1)$ слегка отличаются друг от друга; то же относится и к производным $f'(0)$ и $f'(1)$. Однако отличия эти настолько малы, что не сказываются при реальных оценках, в то время как простота (21), близость формы колебания к обычным кривым блеска цефеид и присутствие гармоник основной частоты весьма удобны.

На рис. 1.1.6 представлен исходный ряд $\{f_k\}$, соответствующий «кривой блеска» (21); добавление к нему гауссовского шума (22) и дало «наблюдаемый» ряд $\{y_k\}$ на рис. 1.1.1. Для того, чтобы характеризовать отношение сигнала к шуму, вычислим дисперсии рядов $\{f_k\}$ и $\{\xi_k\}$, а затем образуем част-

ное

$$S/N \equiv \sigma^2(f)/\sigma^2(\xi). \qquad (1.1.23)$$

В данном случае $\sigma(f) = 0,365$, $\sigma(\xi) = 1/2$, так что $S/N \simeq$ $\simeq 0,53$ - отношение сигнала к шуму достаточно мало. Если для $\{f_k\}$ найти линейную регрессионную зависимость типа (4), то коэффициент α_1 оказывается равным -0,0001, т.е. тренд исходного ряда практически отсутствует, а тренд ряда $\{y_k\}$, описываемый формулами (5), целиком обусловлен флуктуацией случайного компонента $\{\xi_k\}$.

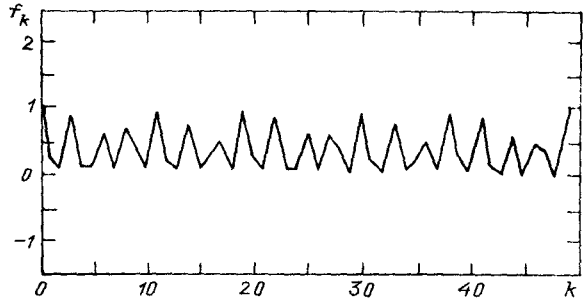


Рис. 1.1.6. Периодический компонент $\{f_k\}$ ряда А

Наибольший интерес представляет, по-видимому, сравнение периодограммы исходного ряда $\{f_k\}$, приведенной на рис. 1.1.7, и периодограммы на рис. 1.1.2 ряда $\{y_k\}$, послужившей основой для выводов о характере процесса. Обнаруженная нами периодическая зависимость действительно единственна, экспериментальное значение ее основного периода $\hat{P}_0 =$ $= 2,6991 = (0,3705)^{-1}$ несколько отличается от истинного периода $P_0 = 2,7183 = (0,3679)^{-1}$. Поскольку колебание (21) не является синусоидальным, периодограмма ряда $\{f_k\}$ помимо основного максимума на частоте $\nu_0 = 0,368$ содержит максимумы на частотах гармоник $2\nu_0 = 0,736$, $3\nu_0 = 1,104$ и т.д. (см. табл. 1.1.1). Между тем, и это весьма существенно для оценки эффективности рассматриваемого метода анализа, гармоники практически не выделяются из шума на периодограмме ряда $\{y_k\}$. Легко понять, что в условиях, когда форма периодической зависимости сильно отличается от синусоидальной (т.е. «энергия» колебания распределяется по большому числу гар-

моник), а отношение сигнала к шуму невелико, чувствительность спектрального метода при выявлении скрытой периодичности снижается, и более предпочтительными оказываются другие методы анализа (см. § 1.7 и гл.4).

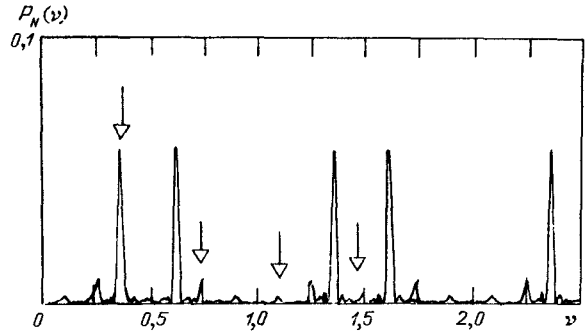


Рис. 1.1.7. Периодограмма $\{f_k\}$ ряда А. Стрелками указаны положения первых четырех гармоник

В отсутствие шума периодограмма ряда $\{f_k\}$ особенно ясно демонстрирует явление подмены частот. Как и следовало ожидать, сопряженные частоты отвечают не только основной гармонике ν_0, но и всем высшим гармоникам $2\nu_0$, $3\nu_0$, ... Поэтому мы можем обобщить выведенную для синусоидального колебания формулу (17) на случай колебания произвольной формы и записать совокупность сопряженных частот в виде

$$\nu_{nl} = \pm n\nu_0 \pm \frac{l}{\Delta t}, \quad n = 1, 2, \ldots, \quad l = 0, 1, \ldots . \quad (1.1.24)$$

На рис. 1.1.8 приведена фазовая диаграмма ряда $\{f_k\}$, свернутого с периодом P_0. Сплошная кривая представляет аппроксимацию (19) тригонометрическим полиномом 4-го порядка, коэффициенты которого были найдены согласно методу наименьших квадратов. Эти значения весьма мало отличаются от точных значений коэффициентов, вычисленных по формуле Фурье (1.4.8) для кривой (21). Точные значения a_n, b_n и величин $A_n^2/4 = (a_n^2 + b_n^2)/4$, характеризующих амплитуды гармоник на периодограмме, даны в пятом - седьмом столбцах табл. 1.1.1. Как сопоставление рисунков 1.1.8 и 1.1.3, так и анализ табл. 1.1.1 показывают, что форма периодического

колебания существенно искажается в присутствии шума при S/N порядка 1.

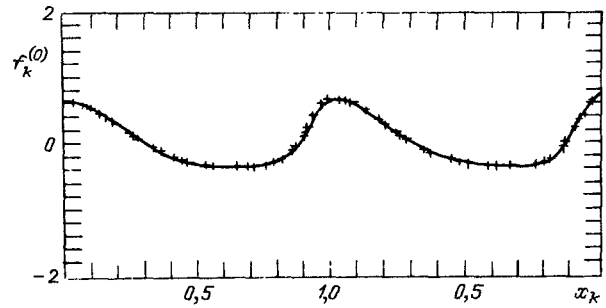


Рис. 1.1.8. Фазовая диаграмма $\{f_k\}$ ряда А для частоты ν=0,3679

Наконец, рис. 1.1.9 иллюстрирует периодограмму чисто стохастического компонента $\xi(t_k)$, а точнее, реализации $\xi(t_k)$, использованной при вычислениях (см. табл. П.7.1). Здесь нужно отметить два следующие обстоятельства.

Во-первых, привлекает внимание сходство периодограмм на рис. 1.1.4 и рис. 1.1.9: периодограмма ряда $\{y_k\}$ после проведения процесса отбеливания колебания с основной частотой ν_0 весьма близка к периодограмме чистого шума. Это позволяет считать, что при аддитивном добавлении шума периодограммы исходного процесса и шума в первом приближении складываются.

Во-вторых, периодограмма шума оказывается весьма изрезанной, причем нередки довольно значительные флуктуации $P_N(\nu)$, указывающие на реальность случайных периодических колебаний заметной амплитуды в шуме. В § 2.3 будет показано, что оба эти свойства отражают существо дела: значения $P_N(\nu)$, разделенные интервалом частот $\delta\nu \simeq (N\Delta t)^{-1}$, оказываются некоррелированными, среднее значение $\langle P_N(\nu)\rangle$ и стандартное отклонение $\sigma[P_N(\nu)]$ для гауссовского белого шума $\xi(t)$ равны, а отношение $P_N(\nu)/\langle P_N(\nu)\rangle$ имеет не гауссовское,

как это иногда думают, а экспоненциальное распределение.

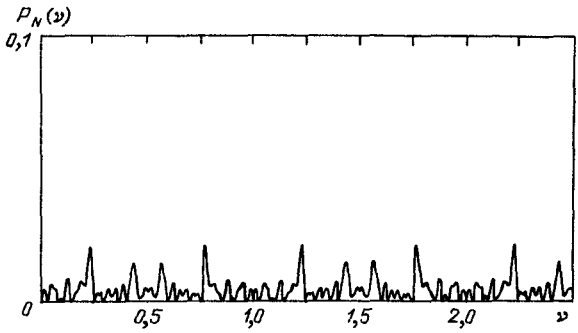


Рис. 1.1.9. Периодограмма стохастического компонеята (ξ_k) ряда A

Характеризуя результаты предварительного анализа ряда A в целом, можно заключить, что в данном случае применение метода Шустера оказалось достаточно эффективным. Существуют и другие методы анализа временных рядов; далее мы рассмотрим для сравнения результаты применения некоторых из них к тому же ряду A. Цель данного параграфа состояла не столько в изложении метода, основанного на вычислении периодограммы Шустера, сколько в описании общих для всех методов анализа понятий.

§ 1.2. Постановка задачи

Вернемся к представлению временного ряда в форме (1.1.2). Заметим прежде всего, что иногда $f_0(t)$ и $f_p(t)$ не разделяют и представляют ряд в виде суммы детерминированного компонента $f(t)$, включая возможность циклического тренда, и стохастического компонента $\xi(t)$:

$$y(t) = f(t) + \xi(t). \qquad (1.2.1)$$

Из дальнейшего обсуждения станет ясно, что принятое в (1.1.2) разделение чаще оказывается более удобным. Например, при изучении переменности блеска короткопериодических цефеид (с продолжительностью цикла порядка 1 часа) функция $f_0(t)$ описывает изменение потока в течение ночи вследствие

изменения высоты объекта над горизонтом, $f_p(t)$ - собственные колебания потока, вносящие основной вклад в переменность $y(t)$, и $\xi(t)$ - шум, обусловленный мерцанием, флуктуациями прозрачности атмосферы, ошибками измерения и прочими случайными факторами. Напротив, основной вклад в наблюдаемые вариации потока излучения от активных галактик вносит стохастический компонент $\xi(t)$, который, разумеется, включает ошибки измерения, но главным образом обусловлен случайными физическими процессами - вспышками в активных ядрах галактик. Соответственно этому информация о процессе находится из анализа стохастического компонента; поскольку при вспышках флуктуации $\xi(t)$ могут превосходить звездную величину, выделение детерминированного слагаемого является непростой задачей (см. § 3.6).

Возможна ситуация, когда временной ряд вообще нельзя разбить на сумму детерминированного и случайного компонентов. Если оставить в стороне тривиальный случай, когда компоненты входят в виде произведения, и стандартное представление (1) получается простым переходом к логарифмам величин, указанная ситуация означает более сложную внутреннюю структуру исходного процесса. Тем не менее реальность представления (1.1.2) следует изучить в первую очередь, и лишь при невозможности достичь удовлетворительного согласия с данными наблюдений целесообразно искать более сложные модели временного процесса. В частности, не исключено сосуществование нескольких периодических колебаний с разными периодами, и кривые блеска цефеид типа TU Cas действительно обнаруживают это явление (Фитч, 1976). Другой пример доставляет рассмотрение так называемых процессов авторегрессии (см. § 2.6), играющих весьма важную роль в прикладном анализе. Более сложные модели временных рядов будут изучены в дальнейшем, а пока сосредоточим внимание на представлениях типа «гармоническое колебание плюс шум».

Принципиальная неоднозначность разделения конкретного временного ряда на отдельные компоненты уже отмечалась в § 1.1. Следует указать и на другое столь же простое, сколь и важное обстоятельство: поскольку наблюдения выделяют *дискретную* совокупность значений исходного процесса, его по-

ведение между моментами наблюдения нельзя восстановить однозначно. Мы можем говорить лишь о наиболее *вероятном* поведении процесса в рамках некоторой модели, согласующейся с доступной информацией о нем.

Условия, при которых исследователь сталкивается с необходимостью объяснить упорядоченную совокупность данных наблюдений, весьма многообразны. Это сильно затрудняет попытки дать полную формулировку задачи анализа временных рядов. Если говорить о достаточно общей формулировке, то «... конечной целью анализа временных рядов ... является достижение более глубокого понимания тех причинных механизмов, которые обусловливают появление этих рядов» (Кендалл и Стюарт, 1966, п. 45.9).

Вместе с тем мы можем перечислить основные этапы, которые включает типичная постановка задачи анализа временных рядов.

I. Освобождение изучаемого ряда от тренда и выпадающих значений. При этом мы не выделяем случай, когда с физической точки зрения интерес представляет именно среднее, систематическое поведение процесса. Обычно выделение тренда и ошибочных значений ряда предваряет другие этапы анализа, однако описание соответствующих процедур целесообразно отложить до тех пор, пока не будет изложена методика обращения с периодическими и случайными рядами. Исходя из этих соображений, этап I описан лишь в гл. 3.

II. Выявление скрытой периодичности данных. Здесь следует обратить внимание на необходимость применять разные методы поиска периодичности, поскольку, как мы увидим далее, эффективность методов существенно зависит от характера периодического процесса, отношения сигнала к шуму и других факторов.

III. Выяснение роли временной скважности данных. Этот этап особенно важен в астрономической практике, которая связана с внесением искусственной периодичности при наземных наблюдениях. Большинство ошибок при интерпретации астрономических наблюдений связано именно с недооценкой временной скважности, которая почти всегда превалирует над случайностью.

IV. Обоснование статистической значимости тех из выделенных детерминированных компонентов - тренда и периодической зависимости, - которые выдержали проверку с точки зрения влияния скважности. Дело в том, что на опыте мы сталкиваемся с конечными реализациями процессов, и потому практически всегда можно найти указания на существование какой-либо детерминированной составляющей, объективно содержащейся в данных наблюдений. Однако от признания реальности явления до обоснования его значимости лежит необходимый этап пути, на котором вступают в дело методы статистического анализа.

V. Изучение стохастического компонента процесса. Не следует путать этот этап с обоснованием статистической значимости. В простейших ситуациях удается показать, что после выделения детерминированных компонентов остается ряд, близкий по характеристикам к белому шуму. Однако обычно стохастическое поведение свойственно самому процессу, и близость соответствующих характеристик к свойствам случайного шума не является обязательной. Мы уже приводили выше пример активности в ядрах галактик. Упомянем в качестве другого примера вспышки звезд типа UV Ceti, где элементы случайности присутствуют как в моментах появления вспышек, так и в их форме, что при адекватном анализе не мешает получить полезную информацию о физической природе процессов, происходящих на поверхности этих звезд. При изучении стохастического компонента цель анализа состоит в построении модели процесса, что позволяет прямо связать наблюдаемые характеристики с физическими явлениями. Так, представление кривой блеска квазара в виде суммы затухающих откликов, обусловленных случайно распределенными во времени вспышками, уже много говорит о природе процессов, ответственных за переменность потока излучения (см. §§ 2.8, 3.6).

Следует несколько подробнее остановиться на характерной черте астрономических наблюдений, упомянутой в предисловии. Речь идет об измерении предельно слабых потоков, что обусловливает в качестве основного режима астрофизических измерений счет отдельных фотонов. Это обстоятельство вносит свою специфику в астрономические временные ряды.

Если среднее время между актами регистрации отдельных фотонов значительно меньше требуемого временного разрешения, то можно оперировать лишь со средними значениями светового потока. При этом достаточно знать, что процесс регистрации фотонов близок к пуассоновскому случайному процессу, и в частности, среднеквадратическая флуктуация числа актов регистрации за время экспозиции равна квадратному корню из среднего числа таких событий за время экспозиции. Однако может случиться так, что характерное время переменности, несущей информацию о физической природе процесса, сравнимо по величине со средним временем между последовательными актами регистрации фотонов. Получающийся при соответствующих наблюдениях ряд отличается от привычных астрономам рядов - он относится к *точечным случайным процессам*. Для анализа таких рядов разработаны специальные методы (см. гл. 5).

Все сказанное выше позволяет отнести анализ временных рядов к классу *обратных задач* математической физики, рассматривающих восстановление свойств исходного процесса по заданному результату действия на него некоторого оператора (это общепринятое определение является, впрочем, настолько широким, что охватывает фактически все задачи интерпретации эмпирических данных). Основная область приложений обратных задач связана с восстановлением изображений, искаженных действием аппаратной функции и шума. Нетрудно заметить, что анализ временных рядов аналогичен восстановлению изображений как методически, так и по самой постановке проблемы. В частности, для стационарных рядов тем оригиналом, который требуется восстановить, является спектральная плотность процесса. Описанный подход рассматривается в гл. 2 и 3.

§ 1.3. Интеграл Фурье

Метод выявления скрытой периодичности, примененный нами в § 1.1, по существу, опирается на спектральный анализ данных. Действительно, формулы (1.1.11) представляют собой *дискретное преобразование Фурье* изучаемого ряда, а периодограмма (1.1.13) - аналог так называемого *спектра мощности* или *спектральной плотности* временного процесса (точный

смысл $P_N(\nu)$ будет рассмотрен в § 1.6). Существуют методы поиска периодичности, не связанные прямым образом со спектральным анализом; так, в гл. 4 мы обсудим методы *непараметрической статистики* и среди них - известный метод Лафлера - - Кинмана. Однако и в этих методах сущность применяемых процедур значительно проясняется, если пользоваться языком спектрального анализа. Вместе с тем понятия спектрального анализа оказались исключительно полезными не только при выявлении скрытой периодичности, но и при изучении общих свойств произвольных временных рядов.

Поскольку имеются превосходные изложения теории интеграла Фурье, и в частности, курс Г.М.Фихтенгольца (1960) и руководство К.Ланцоша (1956), мы ограничимся здесь лишь напоминанием основных понятий, подчеркнув те стороны предмета, которые понадобятся в дальнейшем.

Переходные функции. Предположим, что действительная функция $f(t)$ непрерывна на бесконечном промежутке от $-\infty$ до $+\infty$ и абсолютно интегрируема в нем, т.е.

$$\int_{-\infty}^{\infty} |f(t)|dt < \infty. \qquad (1.3.1)$$

Тогда $f(t)$ может быть представлена в виде *интеграла Фурье*

$$f(t) = \int_{-\infty}^{\infty} e^{i\cdot 2\pi\nu t} F(\nu)d\nu, \qquad i = \sqrt{-1}, \qquad (1.3.2)$$

интерпретируемого как разложение сложной зависимости на простые гармонические колебания всевозможных частот, причем $|F(\nu)|$ представляет собой амплитуду гармоники с частотой ν. Функции, удовлетворяющие условию (1), будем для краткости называть далее *переходными функциями*.

Величина $F(\nu)$ в (2) именуется обычно *спектром* функции $f(t)$. Нахождение спектра опирается на соотношение, обратное (2):

$$F(\nu) = \int_{-\infty}^{\infty} e^{-i\cdot 2\pi\nu t} f(t)dt. \qquad (1.3.3)$$

В общем случае $F(\nu)$ - комплексная функция. Для действительных $f(t)$, которые только и будут рассматриваться далее,

комплексно-сопряженный спектр $F^*(\nu)$ удовлетворяет соотношению

$$F^*(\nu) = F(-\nu), \qquad (1.3.4)$$

так что квадрат модуля спектра

$$G(\nu) \equiv |F(\nu)|^2 = F(\nu)F^*(\nu) \qquad (1.3.5)$$

является симметричной функцией частоты:

$$G(-\nu) = G(\nu). \qquad (1.3.6)$$

Нетрудно проверить, что соотношение (4) обеспечивает действительность $f(t)$ в представлении (2).

При изучении систем постоянно приходится иметь дело с интегралом вида

$$f(t) = \int_{-\infty}^{\infty} f_1(t')f_2(t-t')dt', \qquad (1.3.7)$$

который называется *сверткой* функций $f_1(t)$ и $f_2(t)$. Подставляя (7) в (3), находим после простых преобразований

$$F(\nu) = F_1(\nu)F_2(\nu), \qquad (1.3.8)$$

где $F_1(\nu)$ и $F_2(\nu)$ - спектры функций $f_1(t)$ и $f_2(t)$ соответственно. Таким образом, свертка функций во временно́й области эквивалентна перемножению их спектров в частотной области. Справедливо и утверждение, двойственное только что высказанному: если

$$f(t) = f_1(t)f_2(t), \qquad (1.3.9)$$

то $F(\nu)$ есть свертка спектров функций $f_1(t)$ и $f_2(t)$, а именно

$$F(\nu) = \int_{-\infty}^{\infty} F_1(\nu')F_2(\nu-\nu')d\nu'. \qquad (1.3.10)$$

Введем функцию $\Psi(\tau)$, характеризующую локализацию процесса $f(t)$ во временно́й области:

$$\Psi(\tau) = \int_{-\infty}^{\infty} f(t)f(t+\tau)dt, \quad -\infty < \tau < \infty, \qquad (1.3.11)$$

$$\Psi(-\tau) = \Psi(\tau). \qquad (1.3.12)$$

Как легко видеть на основании (7) и (8), преобразование Фурье этой функции

$$\int_{-\infty}^{\infty} e^{-i\cdot 2\pi\nu\tau}\,\Psi(\tau)d\tau = 2\int_{0}^{\infty}\cos(2\pi\nu\tau)\Psi(\tau)d\tau = \\ = F(\nu)F(-\nu) = G(\nu) \qquad (1.3.13)$$

совпадает с квадратом модуля спектра. Обращение (13) дает представление $\Psi(\tau)$ через $G(\nu)$:

$$\Psi(\tau) = \int_{-\infty}^{\infty} e^{i\cdot 2\pi\nu\tau}\,G(\nu)d\nu = 2\int_{0}^{\infty}\cos(2\pi\nu\tau)G(\nu)d\nu. \qquad (1.3.14)$$

Наконец, полагая $\tau = 0$ в (11) и (14), получаем известное соотношение

$$\int_{-\infty}^{\infty} f^2(t)dt = \int_{-\infty}^{\infty} G(\nu)d\nu, \qquad (1.3.15)$$

составляющее *теорему Рэлея*.

Одну из наиболее обширных областей приложения спектрального анализа представляет радиофизика, где функция $f(t)$ часто описывает изменение тока или напряжения. При этом интеграл в левой части (15) пропорционален полной энергии, выделившейся в течение бесконечно большого промежутка времени, а $G(\nu)$ характеризует распределение этой энергии по частотам. Функцию $G(\nu)$ можно назвать *энергетическим спектром* переходного процесса $f(t)$; не следует смешивать это понятие с рассматриваемым ниже спектром мощности $g(\nu)$ процесса (или спектральной плотностью мощности), описывающим распределение по частотам энергии, выделившейся в единицу времени. Для переходных функций, очевидно, $g(\nu) \equiv 0$.

Приведем примеры спектрального разложения переходных функций.

Пример 1.3.1. Нормальное (гауссово) распределение

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}}\,e^{-t^2/(2\sigma^2)}, \quad -\infty < t < \infty \qquad (1.3.16)$$

характеризуется энергетическим спектром и функцией $\Psi(\tau)$, также заданными формулой Гаусса:

$$G(\nu) = e^{-(\sigma\cdot 2\pi\nu)^2}, \quad -\infty < \nu < \infty, \qquad (1.3.17)$$

$$\Psi(\tau) = \frac{1}{\sigma_0\sqrt{2\pi}} e^{-\tau^2/(2\sigma_0^2)}, \quad -\infty < \tau < \infty, \qquad (1.3.18)$$

где $\sigma_0^2 = 2\sigma^2$.

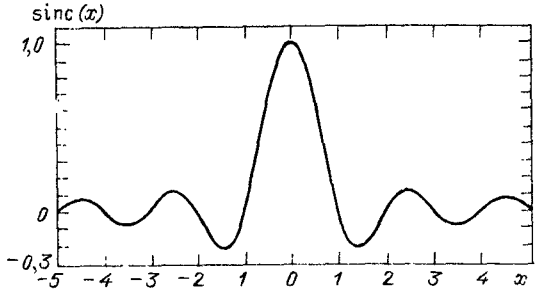


Рис. 1.3.1. Функция sinc(x)

Пример 1.3.2. По-видимому, наиболее часто в теории временных рядов встречается прямоугольное распределение

$$r_T(t) = \begin{cases} 1, & |t| \leqslant T/2, \\ 0, & |t| > T/2. \end{cases} \qquad (1.3.19)$$

Подстановка (19) в формулы (3) и (5) дает

$$R_T(\nu) \equiv \int_{-\infty}^{\infty} e^{-i\cdot 2\pi\nu t} r_T(t)dt = \frac{\sin(\pi\nu T)}{\pi\nu} =$$

$$= T \operatorname{sinc}(\nu T), \quad -\infty < \nu < \infty, \qquad (1.3.20)$$

$$G_T^{(r)}(\nu) = T^2 \operatorname{sinc}^2(\nu T), \qquad (1.3.21)$$

где вспомогательная функция sinc(x) определяется равенством

$$\operatorname{sinc}(x) = \frac{\sin(\pi x)}{\pi x}, \quad -\infty < x < \infty \qquad (1.3.22)$$

(рис. 1.3.1). Свойства функции sinc(x) обсуждаются в Приложении 3; функции $r_T(t)$ и $G_T^{(r)}(\nu)$ представлены на рис. 1.3.2. Расстояние между ближайшими к центру нулями $G_T^{(r)}(\nu)$ равно $2/T$, высота главного максимума (при $\nu = 0$) равна T^2, ближайшие к нему максимумы расположены при $\nu = \pm 2,46/T$, их высота составляет 0,0165 высоты центрального максимума. Функция $\Psi(\tau)$, соответствующая прямоугольному распределению

(19),

$$\Psi_T^{(r)}(\tau) = \begin{cases} T - |\tau|, & |\tau| \leqslant T, \\ 0, & |\tau| > T, \end{cases} \tag{1.3.23}$$

также представлены на рис. 1.3.2.

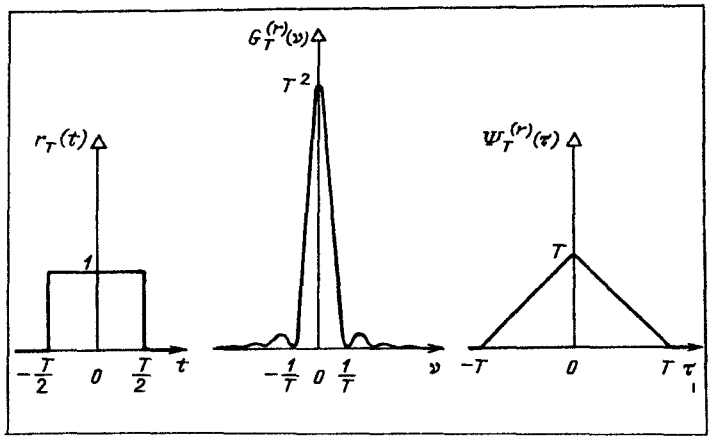


Рис. 1.3.2. Прямоугольное распределение и соответствующие ему функции $G_T(\nu)$ и $\Psi_T(\nu)$

Спектр конечного отрезка функции. Применим свойства свертки к изучению спектров переходных функций, заданных на конечном отрезке времени длиной T. Пусть $f(t)$ определена на промежутке $(-\infty, \infty)$ и ее спектр есть $F(\nu)$; обозначим через $f_T(t)$ отрезок исходной функции

$$f_T(t) = r_T(t) f(t) = \begin{cases} f(t), & |t| \leqslant T/2, \\ 0, & |t| > T/2, \end{cases} \tag{1.3.24}$$

где $r_T(t)$ определена формулой (19). Из (9), (10) и (20) следует:

$$F_T(\nu) \equiv \int_{-\infty}^{\infty} e^{-i\cdot 2\pi\nu t} f_T(t)\,dt =$$

$$= \int_{-\infty}^{\infty} T\ \mathrm{sinc}[(\nu-\nu')T]F(\nu')\,d\nu', \tag{1.3.25}$$

т.е. спектр конечного отрезка функции представляет собой

свертку спектра всей этой функции и sinc (νT). Поскольку ширина главного максимума этой последней функции равна приблизительно $1/T$, соотношение (25) показывает, что все детали истинного спектра процесса $F(\nu)$, имеющие разделение $\leqslant T^{-1}$, будут сглажены в спектре $F_T(\nu)$ конечного отрезка процесса $f_T(t)$. Другими словами, разрешение в спектре отрезка функции длиной T составляет $\delta\nu \simeq T^{-1}$. В дальнейшем мы рассмотрим потерю разрешения при использовании ограниченной информации более подробно.

По мере увеличения длины временного отрезка T первый множитель в подинтегральном выражении (25) стремится, согласно (П. 2.3), к *δ-функции Дирака* (см. Приложение 2). Принимая во внимание формулу замены аргумента (П. 2.5), мы приходим к выводу, что для переходных функций

$$F_T(\nu) \xrightarrow[T \to \infty]{} F(\nu), \quad G_T(\nu) \equiv |F_T(\nu)|^2 \xrightarrow[T \to \infty]{} G(\nu), \qquad (1.3.26)$$

как этого и следовало ожидать из общих соображений. Вновь отметим, что энергетический спектр переходных функций, отнесенный к единице времени, т.е. спектр мощности $g_T(\nu) = G_T(\nu)/T$, стремится с ростом T к нулю.

Пример 1.3.3. Экспоненциально затухающая функция

$$f(t) = \begin{cases} 0, & t < 0, \\ e^{-\gamma t}, & t \geqslant 0, \ \gamma > 0 \end{cases} \qquad (1.3.27)$$

рассматривается на конечном промежутке времени $0 \leqslant t \leqslant T$. Соответствующий энергетический спектр

$$G_T(\nu) = \frac{1}{\gamma^2 + (2\pi\nu)^2}\left[1 + e^{-2\gamma T} - 2e^{-\gamma T}\cos(2\pi\nu T)\right]. \qquad (1.3.28)$$

При $T \longrightarrow \infty$ мы получаем лоренцев профиль $G(\nu)$, встречающийся в классической теории излучения:

$$G(\nu) = \frac{1}{\gamma^2 + (2\pi\nu)^2}, \quad -\infty < \nu < \infty. \qquad (1.3.29)$$

Подстановка этого выражения в (14) дает

$$\Psi(\tau) = \frac{1}{2\gamma} e^{-\gamma|\tau|}, \quad -\infty < \tau < \infty; \qquad (1.3.30)$$

более просто (30) можно найти непосредственно на основании определения (11).

Обобщенный гармонический анализ. Функции, названные выше переходными, введены с целью описать явления типа тренда временных рядов. Однако часто эти явления не укладываются в столь жесткие рамки, какие определяет, скажем, условие (1). Для этого более широкого класса функций интеграл в левой части (15), т.е. полная «энергия» колебания за бесконечно большое время, может расходиться, и, следовательно, функция $\Phi(\tau)$ в (11), вообще говоря, не существует. Тем не менее, может существовать предел

$$\lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} f(t)f(t+\tau)dt \equiv \psi(\tau), \qquad (1.3.31)$$

тождественно равный нулю для переходных функций, а соответствующее преобразование Фурье

$$g(\nu) = \int_{-\infty}^{\infty} e^{-i\cdot 2\pi\nu\tau}\,\psi(\tau)d\tau = 2\int_{0}^{\infty}\cos(2\pi\nu\tau)\psi(\tau)d\tau \qquad (1.3.32)$$

по смыслу представляет собой спектральную плотность детерминированного процесса $f(t)$. Обращение (32) дает

$$\psi(\tau) = \int_{-\infty}^{\infty} e^{i\cdot 2\pi\nu\tau}\,g(\nu)d\nu = 2\int_{0}^{\infty}\cos(2\pi\nu\tau)g(\nu)d\nu. \qquad (1.3.33)$$

Изучение функций, удовлетворяющих условию (31), относится к так называемому *обобщенному гармоническому анализу* (см. Винер, 1933; Давенпорт и Рут, 1958). Имея в виду аналогию с рассматриваемыми ниже периодическими и случайными процессами, $\psi(\tau)$ следует назвать *автоковариационной функцией* детерминированного процесса. Однако для широкого круга практических задач, описание которых удается выполнить в рамках переходных процессов и классического анализа Фурье, интерес представляет функция $\Phi(\tau)$, в то время как $\psi(\tau) \equiv 0$. Напротив, для периодических и стационарных случайных процессов $\Phi(\tau)$ расходится, а $\psi(\tau)$ содержит важную информацию о процессе. Для того, чтобы избежать часто возникающего здесь смешения понятий, мы во всех случаях будем называть $\psi(\tau)$ автоковариационной функцией, а применяемую для описания пе-

реходных процессов функцию $\Psi(\tau)$ оставим без специального названия.

Ограничимся этими краткими замечаниями, касающимися обобщенного гармонического анализа; их достаточно для понимания смысла применяемых далее понятий.

§ 1.4. Периодические функции

Функция $f(t)$ называется *периодической*, если существует интервал P такой, что

$$f(t + P) = f(t) \qquad (1.4.1)$$

для произвольного момента времени t. Очевидно, значения $f(t)$ повторяются и через интервалы, кратные P. Поэтому под основным периодом (или просто периодом) функции P_0 далее будет подразумеваться наименьшее значение P, для которого выполняется условие (1).

Простейшим, и в то же время наиболее важным примером периодической функции служит гармоническое колебание

$$f(t) = A\cos(2\pi\nu_0 t - \varphi_0) =$$
$$= a\cos(2\pi\nu_0 t) + b\sin(2\pi\nu_0 t), \quad -\infty < t < \infty, \qquad (1.4.2)$$

где основная частота

$$\nu_0 = \frac{1}{P_0} \qquad (1.4.3)$$

и коэффициенты a, b связаны с амплитудой A и начальной фазой φ_0 колебания соотношениями

$$a = A\cos\varphi_0, \quad b = A\sin\varphi_0. \qquad (1.4.4)$$

При заданных a, b решение (4) относительно начальной фазы неоднозначно; мы будем использовать это решение в виде

$$A = \sqrt{a^2 + b^2}, \qquad (1.4.5)$$

$$\varphi_0 = \begin{cases} \operatorname{arctg} b/a, & a \geqslant 0,\ b \geqslant 0; \\ \operatorname{arctg} b/a + \pi, & a < 0; \\ \operatorname{arctg} b/a + 2\pi, & a > 0,\ b < 0, \end{cases} \qquad (1.4.6)$$

так что начальная фаза φ_0 (угловое расстояние ближайшего

максимума от начала координат) непрерывно изменяется от 0 до 360°.

Ряд Фурье. Уже пример гармонического колебания показывает, что периодические функции не удовлетворяют основному условию (1.3.1), а именно, они не интегрируемы абсолютно на бесконечном промежутке. Более того, модуль периодической функции не стремится к нулю при $t \to \pm \infty$. Таким образом, периодические функции подчиняются значительно менее жестким ограничениям по сравнению с переходными функциями и, в отличие от последних, не могут быть представлены в виде классического интеграла Фурье.

Можно показать (см., например, Фихтенгольц, 1960, гл. 19), что при весьма слабых ограничениях периодические функции представимы *рядом Фурье*:

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left[a_n \cos(2\pi n\nu_0 t) + b_n \sin(2\pi n\nu_0 t)\right], \qquad (1.4.7)$$

коэффициенты которого - *коэффициенты Фурье* -

$$a_n = \frac{2}{P_0}\int_0^{P_0} \cos(2\pi n\nu_0 t) f(t)dt,$$

$$b_n = \frac{2}{P_0}\int_0^{P_0} \sin(2\pi n\nu_0 t) f(t)dt, \qquad (1.4.8)$$

$n = 0, 1, 2, \ldots$

Обратим внимание, что частоты в (7)

$$\nu_n \equiv n\nu_0 = \frac{n}{P_0}, \qquad n = 1, 2, 3, \ldots, \qquad (1.4.9)$$

кратны основной частоте ν_0, т.е. составляют, как говорят, гармоническую последовательность. При произвольном дискретном наборе ν_n разложение

$$q(t) = \frac{a}{2} + \sum_{n=1}^{\infty}\left[a_n \cos(2\pi\nu_n t) + b_n \sin(2\pi\nu_n t)\right] \qquad (1.4.10)$$

задает так называемую *почти периодическую функцию* $q(t)$, значения которой приближенно повторяются через соответствующие промежутки времени (см. Г. Бор, 1934).

Пользуясь комплексными величинами, ряду Фурье можно

придать более простой вид. Положим

$$c_n \equiv \frac{1}{2}(a_n - ib_n) = \frac{1}{P_0}\int_0^{P_0} e^{-i\cdot 2\pi n\nu_0 t} f(t)dt, \qquad (1.4.11)$$

$$c_{-n} = c_n^*, \quad n = 0, 1, 2, \ldots$$

Тогда из (7) следует искомое представление

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{i\cdot 2\pi n\nu_0 t}, \qquad (1.4.12)$$

напоминающее классическое разложение (1.3.2) по непрерывной области частот. Оба эти представления удается объединить, если воспользоваться δ-функцией Дирака (см. Приложение 2). Именно, формулу (12) можно записать в виде интеграла (1.3.2), где спектр

$$F(\nu) = \sum_{n=-\infty}^{\infty} c_n \delta(\nu - n\nu_0) \qquad (1.4.13)$$

состоит теперь из бесконечного набора равноотстоящих дискретных линий. В частности, спектр гармонического колебания (2) содержит две линии на частотах $\pm\nu_0$:

$$F(\nu) = \frac{A}{2}\left[e^{-i\varphi_0}\,\delta(\nu - \nu_0) + e^{i\varphi_0}\,\delta(\nu + \nu_0)\right]. \qquad (1.4.14)$$

Присутствие в спектре изучаемого процесса резко выделяющихся линий и является обычно признаком наличия в процессе периодической составляющей (см. § 1.6).

Спектр конечного отрезка периодической функции. Подобно тому, как это делалось в § 1.3 для переходных функций, рассмотрим спектр конечного отрезка периодической функции. Для выяснения сути дела достаточно обратиться к отрезку синусоиды

$$f_T(t) = r_T(t)\cdot A\cos(2\pi\nu_0 t - \varphi_0), \qquad (1.4.15)$$

где $r_T(t)$ - прямоугольное распределение (1.3.19). В данном случае условие (1.3.1) выполнено, и мы могли бы найти спектр $F_T(\nu)$ прямым интегрированием согласно (1.3.3). Однако проще воспользоваться свойством свертки (1.3.9), (1.3.10) для преобразования Фурье. Спектр $r_T(t)$ определяется формулой (1.3.20), спектр синусоиды - формулой

(14), так что

$$F_T(\nu) = \frac{A\,T}{2}\left\{ e^{-i\varphi_0}\ \mathrm{sinc}\left[(\nu - \nu_0)\ T\right] + \right.$$
$$\left. + e^{i\varphi_0}\ \mathrm{sinc}\left[(\nu + \nu_0)\ T\right]\right\}, \qquad (1.4.16)$$

где мы приняли во внимание (П.2.5). При $T \to \infty$ спектр $F_T(\nu)$ содержит два выделяющихся пика на частотах $\pm\nu_0$, высота которых пропорциональна T:

$$|F_T(\pm\ \nu_0)| \simeq \frac{AT}{2}\ , \qquad T \to \infty, \qquad (1.4.17)$$

а ширина

$$\delta\nu \simeq \frac{1}{T}\ . \qquad (1.4.18)$$

В соответствии с формулой (П.2.3) по мере увеличения длины отрезка синусоиды его спектр (16) все более приближается к сумме двух δ-функций (14).

Соотношение (18), определяющее ширину линии в спектре гармонического колебания конечной длительности T, заслуживает более подробного рассмотрения. Дело в том, что это соотношение характерно не только для отрезка синусоиды, но и для всякой функции, представляющей ограниченные во времени данные. Именно, можно утверждать, что разрешение в спектре отрезка произвольной функции длины T составляет величину порядка T^{-1}. Для переходных функций это обстоятельство уже отмечалось в связи с формулой (1.3.25), дающей его формальное обоснование. Приведем здесь еще соображения качественного характера.

Вернемся к рассмотрению спектра конечного отрезка синусоиды (15). Ясно, что для хорошей аппроксимации его синусоидами всевозможных частот необходимо взять с максимальным относительным весом (амплитудой) синусоиды той же частоты ν_0. Однако и колебания с близкой к ν_0 частотой войдут в представление, поскольку весьма малое различие частот не проявится в течение конечного промежутка времени T. При какой разнице частот несинхронность заданного и аппроксими-

рующего колебаний становится критической? В качестве граничных значений $\nu_0 \pm \delta\nu/2$ можно принять частоты таких колебаний, которые сначала совпадают по фазе с заданным, а через время T становятся в противофазе ему вследствие различия частот $\delta\nu/2$. Это означает, что набег фазы за время T, составляющий $2\pi(\delta\nu/2)T$, станет равным приблизительно π. Отсюда и следует условие (18). Другими словами, аппроксимирующая синусоида должна совершить за время T не более одного «лишнего» колебания по сравнению с функцией (15).

Ситуация здесь вполне аналогична той, которая обсуждалась в классической теории излучения при нахождении ширины спектральной линии, обусловленной столкновениями. Именно, при достаточно высокой плотности газа можно считать в первом приближении, что лишь столкновения между атомами препятствуют неограниченно долгому процессу излучения каждого из них. В этом последнем случае спектральная линия была бы строго монохроматична; столкновения разбивают бесконечное гармоническое колебание на цуги случайной длины. При этом спектральная линия имеет конечную ширину порядка обратной величины средней длительности одного цуга (см. (5.5.13)).

До сих пор при рассмотрении отрезков периодических функций мы имели дело только с их спектрами $F_T(\nu)$. Обратимся к изучению поведения энергетического спектра

$$G_T(\nu) \equiv |F_T(\nu)|^2, \qquad (1.4.19)$$

т.е. в радиофизической интерпретации, полной энергии вблизи частоты ν, выделившейся за время T. Для отрезка синусоиды (15) имеем из (16) и (19) при $T \to \infty$

$$G_T(\nu) \simeq \frac{A^2T^2}{4}\left\{\mathrm{sinc}^2\left[(\nu-\nu_0)T\right] + \mathrm{sinc}^2\left[(\nu+\nu_0)T\right]\right\}, \qquad (1.4.20)$$

где мы пренебрегли перекрестным членом, малым при значительной длине T. Разумеется, ввиду единообразного поведения периодических функций как при малых, так и при больших значениях t величина $G_T(\nu)$ стремится к бесконечности при $T \to \infty$, т.е. полная выделяющаяся энергия расходится для каждого значения частоты ν. Однако энергия, выделившаяся за единицу

времени, т.е. *спектр мощности*

$$g_T(\nu) = \frac{1}{T}\,|F_T(\nu)|^2, \qquad (1.4.21)$$

остается конечной почти для всех ν. Для отрезка синусоиды из формулы (20) следует

$$g_T(\nu) \simeq \frac{A^2}{4}\left\{T\mathrm{sinc}^2\left[(\nu-\nu_0)T\right] + T\mathrm{sinc}^2\left[(\nu+\nu_0)T\right]\right\}, \qquad (1.4.22)$$

так что, ввиду (П.2.9),

$$g(\nu) = \lim_{T\to\infty} g_T(\nu) = \frac{A^2}{4}\left\{\delta(\nu-\nu_0) + \delta(\nu+\nu_0)\right\}. \qquad (1.4.23)$$

Полученное выражение вполне соответствует по смыслу выражению (14) для спектра синусоиды: частотам $\pm\nu_0$ отвечают гармонические колебания с амплитудами $A/2$, мощность каждого из которых равна $A^2/4$.

Как уже отмечалось в § 1.3, происхождение термина «спектр мощности» связано с изучением временных процессов в радиофизике. Следуя установившейся в последние годы более нейтральной терминологии, мы будем называть $g(\nu)$ *спектральной плотностью* функции $f(t)$.

Для произвольной периодической функции мощность распределена, согласно (13), по дискретной (гармонической) последовательности частот $\nu_n = n\nu_0$, $n=1,2,\ldots$ Поэтому для спектральной плотности периодической функции можно написать по аналогии с (23) следующее выражение:

$$g(\nu) = \sum_{n=-\infty}^{\infty} |c_n|^2\delta(\nu - n\nu_0). \qquad (1.4.24)$$

Автоковариационная функция. Для описания локализации переходного процесса во временной области в предыдущем параграфе была введена функция $\Psi(\tau)$. Если под $f(t)$ понимать периодическую функцию, то интеграл типа (1.3.11) расходится, и это легко понять, так как периодический процесс одинаков на всей временной оси. Для таких процессов по аналогии с (1.3.31) целесообразно определить *автоковариационную функцию* соотношением

$$\psi(\tau) = \frac{1}{P_0}\int_0^{P_0} f(t)f(t+\tau)dt, \quad -\infty < \tau < \infty, \qquad (1.4.25)$$

предусматривающим усреднение произведения $f(t)f(t+\tau)$ по периоду P_0. Подставляя в (25) общее выражение (12) периодической функции в виде ряда, находим:

$$\psi(\tau) = \sum_{n=-\infty}^{\infty} |c_n|^2 e^{i\cdot 2\pi n\nu_0\tau}. \qquad (1.4.26)$$

Как свидетельствуют соотношения (25) и (26), автоковариационная функция периодического процесса является симметричной периодической функцией своего аргумента:

$$\psi(-\tau) = \psi(\tau), \quad \psi(\tau + P_0) = \psi(\tau). \qquad (1.4.27)$$

Нетрудно показать, что $\psi(\tau)$ и $g(\nu)$ связаны для периодических процессов той же парой преобразований Фурье, что и для детерминированных процессов. В самом деле, из (24) и (П.2.5) получаем:

$$\int_{-\infty}^{\infty} e^{i\cdot 2\pi\nu\tau} g(\nu)d\nu = \sum_{n=-\infty}^{\infty} |c_n|^2 \int_{-\infty}^{\infty} e^{i\cdot 2\pi\nu\tau}\ \delta(\nu - n\nu_0)d\nu =$$

$$= \sum_{n=-\infty}^{\infty} |c_n|^2\ e^{i\cdot 2\pi n\nu_0\tau} = \psi(\tau). \qquad (1.4.28)$$

Ввиду несобственного характера δ-функций формальное обращение преобразования Фурье в (28) может вызвать сомнения. Поэтому выполним прямое вычисление интеграла:

$$\int_{-\infty}^{\infty} e^{-i\cdot 2\pi\nu\tau}\ \psi(\tau)d\tau = \sum_{n=-\infty}^{\infty} |c_n|^2 \int_{-\infty}^{\infty} e^{-i\cdot 2\pi(\nu-n\nu_0)\tau}\ d\tau =$$

$$= \sum_{n=-\infty}^{\infty} |c_n|^2\ \delta(\nu - n\nu_0) = g(\nu), \qquad (1.4.29)$$

где мы воспользовались интегральным представлением δ-функции (П.2.4).

Из (25) и (26) при τ=0 следует *теорема Рэлея* для периодических процессов:

$$\frac{1}{P_0}\int_0^{P_0} f^2(t)dt = \sum_{n=-\infty}^{\infty} |c_n|^2 = \int_{-\infty}^{\infty} g(\nu)d\nu. \qquad (1.4.30)$$

Как и ранее, (30) имеет очевидную «энергетическую» интерпретацию.

Спектральная плотность гармонического колебания была

найдена выше; подставляя (23) в (28), получаем автоковариационную функцию гармонического колебания:

$$\psi(\tau) = \frac{A^2}{2}\cos(2\pi\nu_0\tau). \qquad (1.4.31)$$

Это выражение можно вывести и непосредственно из (2) и определения (25).

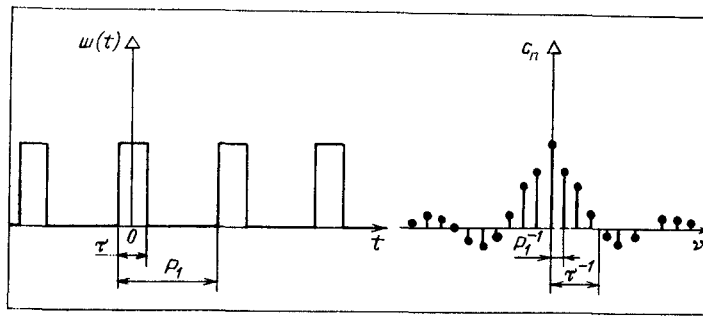


Рис.1.4.1. «Ступенчатая» функция ш(t) и ее амплитудный спектр

Пример 1.4.1. В предыдущем параграфе рассматривалось преобразование Фурье «прямоугольной» функции $r_T(t)$. В дальнейшем нам встретится ее непосредственное обобщение - «ступенчатая» функция ш(t), которую мы зададим следующим образом (см. рис. 1.4.1):

$$\text{ш}(t) = \begin{cases} A, & |t \pm nP_1| \leqslant \tau/2, \quad n = 0, 1, \ldots, \\ 0, & |t \pm nP_1| > \tau/2. \end{cases} \qquad (1.4.32)$$

Найдем разложение периодической функции ш(t) в ряд Фурье. Поскольку значения интегралов (8) не изменяются при сдвиге промежутка интегрирования, имеем:

$$\begin{cases} a_n = \dfrac{2A}{P_1}\displaystyle\int_{-\tau/2}^{\tau/2}\cos\left(2\pi n\,\frac{t}{P_1}\right)dt = \frac{2A}{\pi n}\sin\left(\pi n\,\frac{\tau}{P_1}\right), \\ b_n = \dfrac{2A}{P_1}\displaystyle\int_{-\tau/2}^{\tau/2}\sin\left(2\pi n\,\frac{t}{P_1}\right)dt = 0, \quad n = 0,1,\ldots \end{cases} \qquad (1.4.33)$$

Таким образом, коэффициенты c_n разложения (12) функции ш(t) равны

$$\begin{cases} c_n = A\tau\nu_1 \operatorname{sinc}(\tau\nu_n), \; n = 0, \pm 1, \pm 2, \ldots, \\ \nu_n = n\nu_1 = n/P_1. \end{cases} \qquad (1.4.34)$$

Среднее значение $\langle ш(t)\rangle = c_0 = a_0/2 = A\tau\nu_1$. Отношение длительности «импульса» τ к длине периода P_1 назовем скважностью функции $ш(t)$. Спектр c_n представлен на рис. 1.4.1, непрерывная огибающая имеет вид $c(\nu) = A\tau\nu_1 \operatorname{sinc}(\tau\nu)$. Таким образом, ширина временного окна τ сказывается на эффективной ширине области частот, где амплитуды гармоник не слишком малы.

При известных коэффициентах c_n спектр функции $ш(t)$ определяется формулой (13):

$$Ш(\nu) = \sum_{n=-\infty}^{\infty} c_n \, \delta(\nu - n\nu_1), \qquad (1.4.35)$$

а спектральная плотность и автоковариационная функция – соответственно формулами (24) и (26), где роль основной частоты играет ν_1. Функцию (35) иногда образно называют гребнем Дирака.

§ 1.5. Спектральный анализ стационарных случайных функций

На рис. 1.5.1 сплошной линией представлено изменение отклика аппаратуры $\xi_1(t)$ с течением времени при регистрации потока излучения от звезды. Если бы мы имели возможность вернуться во времени к началу записи и вновь зарегистрировать процесс, две полученные записи $\xi_1(t)$ и $\xi_2(t)$ не были бы идентичными ввиду вероятностного характера процесса излучения фотонов, флуктуаций плотности земной атмосферы и других причин случайного характера. Поэтому реально наблюдавшееся изменение потока излучения звезды $\xi_1(t)$ представляет собой лишь одну из возможных реализаций *случайной функции* или *случайного процесса* $\xi(t)$.

Можно интерпретировать $\xi(t)$ и как обычную случайную величину (см. Приложение 4), которая рассматривается при каждом значении некоторого вспомогательного параметра – в данном случае времени t – и статистические характеристики которой, вообще говоря, различны при отличающихся t. Общее

описание случайного процесса $\xi(t)$ дается совокупностью функций распределения для произвольного числа n и положения «сечений» $t_1, t_2, \ldots, t_n$. Если функции распределения не изменяются при добавлении к $t_1, \ldots, t_n$ одного и того же произвольного числа τ (т.е. инвариантны относительно сдвига вдоль оси времени), то случайный процесс $\xi(t)$ называется *стационарным в узком смысле*. На практике обычно интересуются поведением моментов $\xi(t)$ не выше второй степени, и *стационарным в широком смысле* случайным процессом называется такой процесс, для которого среднее значение постоянно:

$$\langle \xi(t) \rangle = \text{const}, \qquad (1.5.1)$$

а второй момент $\xi(t)$

$$\langle [\xi(t_1) - \langle \xi(t) \rangle][\xi(t_2) - \langle \xi(t) \rangle] \rangle \qquad (1.5.2)$$

зависит лишь от разности $\tau = t_2 - t_1$. Здесь угловые скобки означают усреднение по ансамблю реализаций.

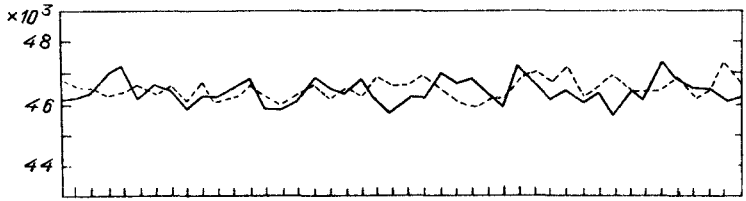


Рис. 1.5.1. Реализации стационарного случайного процесса

Автоковариационная функция и спектральная плотность. Пусть $\xi(t)$ - действительный стационарный в широком смысле случайный процесс, который мы будем полагать для простоты центрированным:

$$\langle \xi(t) \rangle = 0. \qquad (1.5.3)$$

Назовем

$$\psi(\tau) = \langle \xi(t)\, \xi(t + \tau) \rangle \qquad (1.5.4)$$

автоковариационной функцией процесса. Очевидно, $\psi(\tau)$ действительна, симметрична, а ее значение при $\tau = 0$ есть дисперсия случайного процесса:

$$\psi(0) = \langle \xi^2(t) \rangle \equiv \sigma^2. \qquad (1.5.5)$$

Наша ближайшая задача - ввести для процесса $\xi(t)$ понятие спектральной плотности $g(\nu)$, описывающей интенсивность монохроматических случайных составляющих $\xi(t)$.

Рассмотрим с этой целью конечный отрезок $\xi_T(t)$ одной из реализаций процесса $\xi(t)$, заданного на всей временной оси. Именно, положим

$$\xi_T(t) = r_T(t)\ \xi(t) = \begin{cases} \xi(t), & |t| \leqslant T/2, \\ 0, & |t| > T/2. \end{cases} \qquad (1.5.6)$$

Пусть

$$F_T(\nu) = \int_{-\infty}^{\infty} e^{-i\cdot 2\pi\nu t}\ \xi_T(t)dt = \int_{-T/2}^{T/2} e^{-i\cdot 2\pi\nu t}\ \xi(t)dt \qquad (1.5.7)$$

- функция типа (1.3.3), но теперь являющаяся случайной величиной для каждого значения ν, т.е. также представляющая собой случайный процесс. Среднее значение

$$\langle F_T(\nu)\rangle = 0; \qquad (1.5.8)$$

найдем среднее квадрата модуля $F_T(\nu)$, т.е.

$$\langle |F_T(\nu)|^2\rangle = \langle F_T(\nu)\ F_T^*(\nu)\rangle =$$

$$= \int_{-T/2}^{T/2} dt' \int_{-T/2}^{T/2} e^{-i\cdot 2\pi\nu(t'-t'')} \langle \xi(t')\ \xi(t'')\rangle dt''. \qquad (1.5.9)$$

Воспользовавшись определением (4) и изменив порядок интегрирования, нетрудно получить

$$\langle |F_T(\nu)|^2\rangle = 2\int_0^T \cos(2\pi\nu\tau)(T-\tau)\psi(\tau)d\tau. \qquad (1.5.10)$$

Для переходных функций $|F_T(\nu)|^2$ стремится при $T\to\infty$ к энергетическому спектру $G(\nu)$, однако если пытаться перейти к пределу в соотношении (10), то получится расходимость, пропорциональная T. Причина этого заключается в том, что реализации случайного процесса не обязаны убывать на бесконечности как переходные функции; реализации «в среднем одинаковы» на всей временной оси. Поэтому представляется целесо-

образным сначала разделить (10) на T и ввести

$$g_T(\nu) = \frac{1}{T} <| F_T(\nu) |^2> \qquad (1.5.11)$$

и лишь затем устремить T к бесконечности:

$$g(\nu) = \lim_{T\to\infty} g_T(\nu) = 2\int_0^\infty \cos(2\pi\nu\tau)\psi(\tau)d\tau. \qquad (1.5.12)$$

Функция $g(\nu)$, определенная соотношениями (11) и (12), называется *спектральной плотностью случайного процесса* $\xi(t)$. Из (12) видно, что $g(\nu)$ - симметричная функция частоты. Эту формулу, а также обратную ей формулу, выражающую автоковариационную функцию через спектральную плотность, можно представить в стандартном виде:

$$g(\nu) = \int_{-\infty}^{\infty} e^{-i\cdot 2\pi\nu\tau}\psi(\tau)d\tau, \qquad (1.5.13)$$

$$\psi(\tau) = \int_{-\infty}^{\infty} e^{i\cdot 2\pi\nu\tau} g(\nu)d\nu = 2\int_0^\infty \cos(2\pi\nu\tau)g(\nu)d\nu$$

(*теорема Винера-Хинчина*). Полагая в последнем соотношении $\tau = 0$ и вспоминая (5), находим также аналог теоремы Рэлея для случайных функций:

$$\sigma^2 = \int_{-\infty}^{\infty} g(\nu)d\nu. \qquad (1.5.14)$$

Формула (14) позволяет трактовать $g(\nu)d\nu$ как дисперсию тех гармонических компонентов $\xi(t)$, частоты которых заключены в бесконечно узком интервале $[\nu, \nu + d\nu]$.

Нетрудно заметить близкое сходство трактовки многих понятий для регулярных, периодических и случайных процессов; это относится, в частности, к определениям автоковариационной функции и спектральной плотности. Однако $\psi(\tau)$ случайного процесса не обязательно должна быть периодична, как автоковариационная функция периодического процесса, а его спектральная плотность $g(\nu)$ соответственно не обязана содержать лишь дискретный набор линий, хотя такая возможность не исключена (см. ниже пример 1.5.2).

Белый шум. Рассмотрим стационарный случайный процесс, называемый белым шумом. Именно, предположим, что значения

случайной функции $\xi(t)$, разделенные сколь угодно малым промежутком времени $\tau \neq 0$, статистически не зависят друг от друга, так что $\psi(\tau) = 0$ при $\tau \neq 0$. Формально это обстоятельство можно выразить, воспользовавшись δ-функцией Дирака:

$$\psi(\tau) = g_0\delta(\tau), \quad g_0 = \text{const}. \tag{1.5.15}$$

Подставляя (15) в (13) и принимая во внимание формулу (П.2.5), получаем:

$$g(\nu) = g_0, \quad -\infty < \nu < \infty, \tag{1.5.16}$$

- спектральная плотность белого шума постоянна на всем бесконечном промежутке частот. Понятие белого шума является идеализацией поведения физических процессов, для которых $g(\nu)$ сохраняет приблизительно постоянное значение в широком диапазоне частот, т.е. интенсивность быстрых колебаний случайного процесса сравнима с интенсивностью медленных изменений.

Определение белого шума не касается вида функции распределения $\xi(t)$; в частности, если плотность распределения имеет вид

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-x^2/(2\sigma^2)}, \quad -\infty < x < \infty, \tag{1.5.17}$$

случайный процесс называют *нормальным* или *гауссовским белым шумом*. Следует также заметить, что белый шум может отвечать не только независимым, но лишь некоррелированным между собой значениям процесса в различные моменты времени.

Спектральное разложение стационарного процесса. Мы пришли выше к понятию спектральной плотности случайного процесса $\xi(t)$, исходя из рассмотрения спектра конечного отрезка процесса $\xi_T(t)$ и переходя затем к пределу $T \to \infty$. Этот наглядный путь можно значительно сократить, если воспользоваться δ-функцией Дирака (Приложение 2) и ввести формально случайный спектр $F(\nu)$ всего процесса $\xi(t)$ таким же образом, как это было сделано в (1.3.2) для детерминированных про-

цессов:

$$\xi(t) = \int_{-\infty}^{\infty} e^{i\cdot 2\pi\nu t}\, F(\nu)\,d\nu. \qquad (1.5.18)$$

Подставляя это выражение в (4) и учитывая вещественность $\xi(t)$, получаем

$$\psi(\tau) = \iint_{-\infty}^{\infty} e^{-i\cdot 2\pi\nu t + i\cdot 2\pi\nu'(t+\tau)} \langle F^{*}(\nu)F(\nu')\rangle d\nu d\nu'. \qquad (1.5.19)$$

Отсюда видно, что единственный способ обеспечить зависимость автоковариационной функции только от τ, как это и должно быть для стационарных процессов, заключается в требовании принять среднее $\langle F^{*}(\nu)F(\nu')\rangle$ равным нулю везде, кроме «диагонали» в плоскости (ν, ν'):

$$\langle F^{*}(\nu)\, F(\nu')\rangle = g(\nu)\, \delta(\nu - \nu'). \qquad (1.5.20)$$

Тогда ввиду (П.2.5) мы вновь приходим ко второму из соотношений (13) и, следовательно, должны интерпретировать $g(\nu)$ в (20) как спектральную плотность случайного процесса.

Случайные процессы с дискретным спектром. Хорошей иллюстрацией приведенных выше общих положений служит изучение процессов с дискретным спектром, называемых также *периодическими стационарными случайными процессами.* Обращение к ним тем более полезно, что оно расширяет понятие о периодических процессах и необходимо при интерпретации результатов, касающихся поисков периодических компонентов в экспериментальных данных.

Рассмотрим сначала простейший процесс указанного типа, а именно

$$\xi(t) = A\cos(2\pi\nu_0 t - \varphi), \qquad (1.5.21)$$

где амплитуда A и фаза φ суть независимые случайные величины с плотностями распределения $f_A(x)$, $-\infty < x < \infty$, и $f_\varphi(y) = 1/(2\pi)$, $0 \leqslant y \leqslant 2\pi$, соответственно. Процесс (21) можно представить ансамблем синусоид с частотой ν_0, сдвинутых друг относительно друга по фазе «совершенно случайным образом», амплитуды которых выбираются в соответствии с некоторым законом $f_A(x)$. Нетрудно убедиться, что при произвольном

распределении амплитуд $\langle \xi(t) \rangle = 0$ и автоковариационная функция

$$\psi(\tau) = \frac{\langle A^2 \rangle}{2} \cos(2\pi \nu_0 \tau). \qquad (1.5.22)$$

Сравнение (22) с автоковариационной функцией детерминированного гармонического колебания (1.4.31) показывает, что рандомизация фазы не сказалась на виде $\psi(\tau)$. Из (13) и (22) находим также выражение для спектральной плотности процесса (21):

$$g(\nu) = \frac{\langle A^2 \rangle}{4} \left[\delta(\nu - \nu_0) + \delta(\nu + \nu_0) \right], \qquad (1.5.23)$$

которое отличается от (1.4.23) лишь дополнительным усреднением квадрата амплитуды.

Перейдем теперь к общему случаю действительного процесса с дискретным спектром $\xi(t)$, определив его условием периодичности автоковариационной функции

$$\psi(\tau + P_0) = \psi(\tau), \quad -\infty < \tau < \infty, \qquad (1.5.24)$$

которому удовлетворяет и частный случай (22). Заметим, что $\xi(t)$ предполагается для простоты центрированной, и мы используем, как обычно, обозначение $P_0 = \nu_0^{-1}$.

Поскольку всякую периодическую функцию можно разложить в ряд Фурье, мы имеем согласно (1.4.11) и (1.4.12):

$$\psi(\tau) = \sum_{n=-\infty}^{\infty} g_n \, e^{i \cdot 2\pi n \nu_0 \tau}, \qquad (1.5.25)$$

$$g_n = \frac{1}{P_0} \int_0^{P_0} e^{-i \cdot 2\pi n \nu_0 \tau} \, \psi(\tau) d\tau. \qquad (1.5.26)$$

Легко показать, что ввиду (24) выполняется равенство

$$\langle | \xi(t) - \xi(t + P_0) |^2 \rangle = 0, \qquad (1.5.27)$$

которое свидетельствует о периодичности почти всех выборочных функций процесса $\xi(t)$ с периодом P_0. Но тогда $\xi(t)$ так-

же можно разложить в ряд Фурье со случайными коэффициентами

$$\begin{cases} \xi(t) = \sum_{k=-\infty}^{\infty} \xi_k e^{i\cdot 2\pi k\nu_0 t}, & (1.5.28) \\ \xi_k = \frac{1}{P_0}\int_0^{P_0} e^{-i\cdot 2\pi k\nu_0 t}\,\xi(t)dt, \quad k = 0,1,\ldots, & (1.5.29) \end{cases}$$

причем $\xi_0 = 0$ и при различных k и l коэффициенты ξ_k, ξ_l не коррелированы:

$$\langle \xi_k \xi_l^* \rangle = g_k \delta_{k,l}. \qquad (1.5.30)$$

Спектральная плотность $g(\nu)$ периодического стационарного случайного процесса находится путем подстановки (25) в первую из формул (13):

$$g(\nu) = \sum_{n=-\infty}^{\infty} g_n\,\delta(\nu - n\nu_0) = \sum_{n=-\infty}^{\infty} \langle |\xi_n|^2 \rangle\,\delta(\nu - n\nu_0). \qquad (1.5.31)$$

Дискретные спектральные линии здесь эквидистантны, поскольку процесс состоит из совокупности случайно сдвинутых по фазе периодических функций. Конкретным представлением общих формул (18)-(20) служат в данном случае формулы (28), (25) и (30) соответственно.

Точное соотношение между $g_T(\nu)$ и $g(\nu)$. Формула (12) устанавливает тот важный факт, что спектральная плотность конечного отрезка процесса $g_T(\nu)$ стремится при увеличении длины промежутка времени T к истинной спектральной плотности $g(\nu)$. Однако нередко приходится выяснять, каким именно образом происходит предельный переход; для этого нужно найти соотношение между $g_T(\nu)$ и $g(\nu)$, справедливое при любых конечных значениях T.

Такое соотношение нетрудно получить, записав (10) и (11) в виде

$$g_T(\nu) = \int_{-T}^{T} \cos(2\pi\nu\tau)\left(1 - \frac{|\tau|}{T}\right)\psi(\tau)d\tau \qquad (1.5.32)$$

и подставив сюда выражение (13) автоковариационной функции через $g(\nu)$. После элементарных преобразований находим

искомое соотношение

$$g_T(\nu) = T \int_{-\infty}^{\infty} \mathrm{sinc}^2\left[(\nu - \nu')T\right] g(\nu')\,d\nu', \qquad (1.5.33)$$

где $\mathrm{sinc}(x) = \sin(\pi x)/(\pi x)$. Согласно (П.2.9) при $T \to \infty$ мы имеем $T\,\mathrm{sinc}^2(\nu T) \to \delta(\nu)$, так что ядро интегрального представления (33) становится все более «узким», а $g_T(\nu)$ стремится к $g(\nu)$.

Линейная фильтрация случайных процессов. Рассмотрим стационарный случайный процесс $\xi(t)$ со средним $\langle\xi(t)\rangle = 0$, автоковариационной функцией $\psi_\xi(\tau)$ и связанной с ней посредством (13) спектральной плотностью $g_\xi(\nu)$. Пусть далее $h(\tau)$ - некоторая детерминированная функция с «хорошим» поведением при $\tau \to \pm\infty$, как это свойственно переходным функциям. Случайному процессу $\xi(t)$ можно сопоставить новый случайный процесс $\eta(t)$, связанный с исходным соотношением

$$\eta(t) = \int_{-\infty}^{\infty} h(t - t')\xi(t')\,dt'. \qquad (1.5.34)$$

В радиофизике t обычно означает текущий момент времени, и (34) описывает действие линейного фильтра на входной сигнал $\xi(t)$. Более общий подход включает добавление шума, но сейчас нас интересует лишь действие линейного фильтра. Поскольку данное состояние системы зависит только от ее прошлого поведения, ядро $h(\tau)$ - *импульсная характеристика* - считается при указанной интерпретации равным нулю для $\tau<0$.

Если параметр t представляет собой пространственную координату, скажем, отсчет вдоль записи спектра звезды, то упомянутое условие не является обязательным. Преобразование (34) можно тогда трактовать как сглаживание исходной записи $\xi(t)$ при помощи весовой функции $h(\tau)$.

Стоящая перед нами задача заключается в нахождении статистических свойств прошедшего через фильтр сигнала $\eta(t)$. Очевидно, его среднее значение $\langle\eta(t)\rangle = 0$. Ав-

токовариационная функция

$$\psi_\eta(\tau)\equiv\langle\eta(t)\eta(t+\tau)\rangle = \int_{-\infty}^{\infty} h(t')dt' \int_{-\infty}^{\infty} h(t'')\psi_\xi(\tau+t'-t'')dt''. \qquad (1.5.35)$$

Вводя

$$h_1(\tau) = \int_{-\infty}^{\infty} h(t')\ h(t'+\tau)dt', \qquad (1.5.36)$$

запишем соотношение (35) следующим образом:

$$\psi_\eta(\tau) = \int_{-\infty}^{\infty} h_1(\tau - \tau')\psi_\xi(\tau')d\tau'. \qquad (1.5.37)$$

Мы видим, что автоковариационная функция также подвергается линейной фильтрации, но ядро фильтра квадратично зависит от ядра $h(\tau)$.

Для того, чтобы найти спектральную плотность $g_\eta(\nu)$ выходного сигнала, применим к (37) преобразование Фурье. По теореме о свертке

$$g_\eta(\nu) = H_1(\nu)g_\xi(\nu), \qquad (1.5.36)$$

где

$$H_1(\nu) = \int_{-\infty}^{\infty} e^{-i\cdot 2\pi\nu\tau}\ h_1(\tau)d\tau = |H(\nu)|^2, \qquad (1.5.39)$$

$$H(\nu) = \int_{-\infty}^{\infty} e^{-i\cdot 2\pi\nu\tau}\ h(\tau)d\tau. \qquad (1.5.40)$$

Преобразование Фурье $H(\nu)$ импульсной характеристики $h(\tau)$ называется в радиофизике *передаточной функцией.* Это название широко распространено и в оптике, где $h(\tau)$ часто именуют *функцией рассеяния точки* или, как в математике, *функцией Грина.* Основной результат теории линейной фильтрации заключается в том, что фильтру с импульсной характеристикой $h(\tau)$ соответствует изменение спектральной плотности, сводящееся к умножению исходной плотности на квадрат модуля передаточной функции:

$$g_\eta(\nu) = |H(\nu)|^2 g_\xi(\nu). \qquad (1.5.41)$$

Целью фильтрации является обычно повышение отношения сигнала к шуму. Приведем простые примеры, описывающие типичные на практике ситуации.

Пусть изучается оптическое изображение, полученное при помощи системы с разрешением порядка Δx. В области пространственных частот это соответствует распределению сигнала вплоть до частот порядка $(\Delta x)^{-1}$, более высокие частоты отвечают случайному шуму типа зернистости фотоматериала. Подбирая верхнюю границу ν_{max} фильтра низких частот так, чтобы $\nu_{max} \approx (\Delta x)^{-1}$, мы добьемся более уверенного выделения изображения на фоне шума. Ясно, что имеет смысл говорить о нахождении оптимального (в некотором принятом понимании) фильтра, и такие оптимальные фильтры действительно существуют.

Второй пример связан с методикой астрономических наблюдений в инфракрасной области спектра. Ввиду высокого, а главное - непостоянного уровня шума, обусловленного тепловым свечением неба, телескопа и самого приемника, наблюдатель достаточно часто (скажем, с частотой $\nu_0 \simeq 20$ Гц) попеременно наводится на исследуемый объект и участок неба вблизи него. При этом сигнал от объекта модулируется с частотой ν_0, а дальнейшее его выделение опирается на использование узкополосного фильтра, центрированного на частоту ν_0.

Наконец, третий пример отвечает простому усреднению исходного процесса $\xi(t)$ с весовой функцией

$$h(\tau) = \begin{cases} T^{-1}, & -T \leqslant \tau \leqslant 0, \\ 0, & \tau < -T,\ \tau > 0, \end{cases} \qquad (1.5.42)$$

так что

$$\eta(t,T) = \frac{1}{T}\int_t^{t+T} \xi(t')dt'. \qquad (1.5.43)$$

Мы имеем в данном случае $|H(\nu)|^2 = \mathrm{sinc}^2(\nu T)$ и

$$g_\eta(\nu,T) = \mathrm{sinc}^2(\nu T)\ g_\xi(\nu). \qquad (1.5.44)$$

Частоты, пропускаемые рассматриваемым фильтром, занимают область приблизительно вплоть до $|\nu_{max}| \simeq T^{-1}$ - усреднение сглаживает резкие колебания. Еще более значительного сужения эффективной полосы частот можно достичь при использовании n-кратного последовательного усреднения процесса (тогда в (44) вместо показателя степени 2 появится $2n$).

Формула (44) позволяет качественно понять смысл требований, обеспечивающих *эргодичность* случайного процесса (см. Ахманов и др., 1981). Фиксируем в (43) начальный момент времени t и будем рассматривать свойства $\eta(t,T)$ по мере увеличения продолжительности интервала усреднения T. Как известно, процесс называется эргодическим, если для него временное среднее $\eta(t,T)$ стремится с ростом T к статистическому среднему $<\xi(t)>$. Очевидно, что для эргодического процесса дисперсия $\sigma_\eta^2(T)$ величины (43) должна стремиться при $T \to \infty$ к нулю. Из (44) и теоремы Рэлея следует:

$$\sigma_\eta^2(T) = \int_{-\infty}^{\infty} g_\eta(\nu,T)\,d\nu = \int_{-\infty}^{\infty} \mathrm{sinc}^2(\nu T)\, g_\xi(\nu)\,d\nu, \qquad (1.5.45)$$

так что с ростом T интеграл в (45) охватывает фактически лишь низкочастотную область $|\nu| \leqslant T^{-1}$, и поведение $\sigma_\eta^2(T)$ при $T \to \infty$ определяется поведением $g_\xi(\nu)$ при $\nu \to 0$. Так, если спектральная плотность процесса $\xi(t)$ конечна в нуле, то согласно (45) и (П.3.4)

$$\sigma_\eta^2(T) \simeq g_\xi(0)\int_{-\infty}^{\infty} \mathrm{sinc}^2(\nu T)\,d\nu = g_\xi(0)\,\frac{1}{T} \qquad (1.5.46)$$

- процесс $\xi(t)$ эргодичен. Напротив, усреднение по времени так называемого *фликкер-шума*, характеризующегося спектральной плотностью $g_\xi(\nu) \propto \nu^{-1}$, приводит к дисперсии $\sigma_\eta^2(T) =$ = const, т.е. фликкер-шум не обладает свойством эргодичности.

§ 1.6. Реальные процессы

Временные процессы, с которыми приходится сталкиваться на практике, лишь в исключительных случаях могут рассматриваться как чисто детерминированные, либо чисто стохастические. Чаще всего процессы включают оба компонента, по-

разному проявляющиеся в данных наблюдений конечной длительности. Целью данного параграфа является рассмотрение соотношения между отдельными компонентами временного процесса на примере простейшей и наиболее важной модели (1.1.2), когда эти компоненты входят аддитивным образом.

Связь между периодограммой и спектральной плотностью. Итак, предположим, что изучаемый процесс допускает разбиение на тренд, периодическую зависимость и стационарный случайный компонент в форме (1.1.2). Тогда спектр части процесса длительности T можно записать следующим образом:

$$F_T^{(y)}(\nu) = F_T^{(0)}(\nu) + F_T^{(p)}(\nu) + F_T^{(\xi)}(\nu), \qquad (1.6.1)$$

где слагаемые отвечают перечисленным компонентам. Найдем среднее (по ансамблю) значение квадрата модуля суммарного спектра $F_T^{(y)}(\nu)$. Если пренебречь перекрестными членами, ожидаемый вклад которых мал ввиду несогласованного поведения отдельных слагаемых в (1), то

$$\langle |F_T^{(y)}(\nu)|^2 \rangle \simeq |F_T^{(0)}(\nu)|^2 + |F_T^{(p)}(\nu)|^2 + \langle |F_T^{(\xi)}(\nu)|^2 \rangle. \qquad (1.6.2)$$

Учитывая определение энергетического спектра (1.3.26), а также определение спектральной плотности периодических и случайных процессов (1.4.21) и (1.5.11), перепишем последнее соотношение в виде

$$\frac{1}{T} \langle |F_T^{(y)}(\nu)|^2 \rangle \simeq \frac{1}{T} G_T(\nu) + g_T^{(p)}(\nu) + g_T^{(\xi)}(\nu). \qquad (1.6.3)$$

В случае, когда тренд описывается переходным процессом, первое слагаемое в (3) стремится к нулю с ростом T; в более общем случае оно сходится к спектральной плотности детерминированного процесса, определенной в смысле (1.3.31) и (1.3.32). Поэтому

$$\frac{1}{T} \langle |F_T^{(y)}(\nu)|^2 \rangle \xrightarrow[T\to\infty]{} g^{(0)}(\nu) + g^{(p)}(\nu) + g^{(\xi)}(\nu), \qquad (1.6.4)$$

и при больших значениях T величину $T^{-1}\langle |F_T^{(y)}(\nu)|^2 \rangle$ можно рассматривать в качестве оценки суммарной спектральной плотности процесса $y(t)$.

Отметим некоторые особенности, существенные для практических приложений соотношения (3).

Прежде всего, как свидетельствует (1.4.22), спектральная плотность периодического процесса на частотах гармоник характеризуется асимптотическим поведением

$$g_T^{(p)}(\nu) \simeq \frac{A^2}{4} T, \quad T \to \infty, \tag{1.6.5}$$

т.е. высота линий неограниченно увеличивается с ростом T. По этой причине удобнее рассматривать

$$P_T(\nu) \equiv \frac{1}{T^2} |F_T^{(y)}(\nu)|^2, \tag{1.6.6}$$

так что, согласно (3),

$$\langle P_T(\nu) \rangle \simeq \frac{1}{T} \left[g_T^{(0)}(\nu) + g_T^{(p)}(\nu) + g_T^{(\xi)}(\nu) \right], \tag{1.6.7}$$

и периодическому компоненту будут соответствовать спектральные линии, высота которых практически не изменяется с увеличением длины промежутка наблюдений T. При этом вклад систематического компонента, если он описывается переходным процессом, будет убывать пропорционально T^{-2}, а вклад стохастического компонента - пропорционально T^{-1} (Шустер, 1906; Диминг, 1975). Вместе с тем следует иметь в виду, что тренд далеко не всегда удается представить переходной функцией, он может иметь и отличную от нуля спектральную плотность. Тогда вклад систематического компонента в $\langle P_T(\nu) \rangle$ убывает по мере роста T как T^{-1}. С другой стороны, стохастический компонент может содержать случайное периодическое колебание типа (1.5.21) с подходящей частотой, и тогда трактовка случайного колебания практически не отличается от трактовки детерминированного периодического процесса - оба компонента вносят вклад, не зависящий от T.

В любом случае увеличение длины реализации процесса приводит к повышению контраста дискретных линий, отвечающих периодическим компонентам.

При чисто синусоидальной форме колебания с аплитудой A высота линий, как видно из (5) и (7), равна

$$\langle P_T(\nu_0)\rangle \simeq A^2/4, \quad T\to\infty \qquad (1.6.8)$$

(Шустер, 1906). Если же нас интересует высота линии, соответствующей n-ой гармонике разложения Фурье (1.4.7) произвольной периодической функции, то ввиду аналогичных соображений

$$\langle P_T(\nu_n)\rangle \simeq A_n^2/4 = \frac{a_n^2 + b_n^2}{4}, \quad T\to\infty. \qquad (1.6.9)$$

Обратим внимание в связи с (9) на соответствие данных последнего столбца табл. 1.1.1 и расчетов периодограммы $\{f_k\}$, представленной на рис. 1.1.7.

Строго говоря, достаточно надежно оценить амплитуду периодического колебания на основании (8) удается лишь в том случае, если вблизи ν_0 малы спектральные плотности шума и регулярного слагаемого (см. (7)). Информация же о фазе колебания почти полностью теряется при построении периодограммы конечного отрезка временного ряда. Для восстановления этой информации можно воспользоваться соотношением (1.1.12) при частоте $\nu = \nu_0$, соответствующей найденной с помощью периодограммы основной частоте периодического компонента. Другой, более надежный способ основан на построении фазовой диаграммы, как это делалось в § 1.1.

Второе замечание касается операции статистического усреднения, которую мы применили в формулах (2) - (4). На практике исследователь имеет в своем распоряжении лишь один образец изучаемого процесса и, разумеется, не может произвести статистическое усреднение по ансамблю. Непрерывные наблюдения позволяют найти лишь одну из возможных реализаций (6), т.е. некоторую случайную функцию частоты. При конечной длительности промежутка наблюдения (6) может обладать особенностями (в частности, содержать максимумы), которые не присущи истинным спектральным плотностям (7). Это обстоятельство требует разработки соответствующей статистической теории для оценки значимости флуктуаций $P_T(\nu)$. Менее очевиден следующий весьма важный факт: выборочная, т.е. отвечающая какой-либо реализации случайного компонен-

та, функция $TP_T^{(\xi)}(\nu)$ не стремится при $T \to \infty$ к истинной спектральной плотности $g^{(\xi)}(\nu)$ этого компонента ни в обычном смысле, ни в смысле статистической сходимости. Как говорят статистики, величина $TP_T^{(\xi)}(\nu)$ не может считаться состоятельной оценкой спектральной плотности; в физической литературе более привычной была бы ссылка на отсутствие свойства эргодичности. Обсуждение относящихся сюда вопросов приводится в §§ 2.3 и 2.4.

Дискретность данных наблюдений является не менее важным фактором, чем их уникальность и конечный объем. В самом деле, на практике измерениям доступны значения функции $y(t)$ лишь для дискретной совокупности моментов времени t_0, t_1, ..., t_{N-1}, которые мы будем считать сейчас не обязательно равноотстоящими друг от друга. Поэтому вместо (6) вычисляют величину

$$P_N(\nu) \equiv \frac{1}{N^2} \left| F_N^{(y)}(\nu) \right|^2, \qquad (1.6.10)$$

где

$$F_N^{(y)}(\nu) = \sum_{k=0}^{N-1} e^{-i \cdot 2\pi\nu t_k} y^{(0)}(t_k). \qquad (1.6.11)$$

Две последние формулы можно записать также в виде:

$$\begin{aligned} P_N(\nu) &= U_N^2(\nu) + V_N^2(\nu), \\ U_N(\nu) &= \frac{1}{N} \sum_{k=0}^{N-1} \cos(2\pi\nu t_k) y_k^{(0)}, \\ V_N(\nu) &= \frac{1}{N} \sum_{k=0}^{N-1} \sin(2\pi\nu t_k) y_k^{(0)}. \end{aligned} \qquad (1.6.12)$$

Мы вновь пришли к периодограмме процесса $P_N(\nu)$, которая уже использовалась при поисках периодичности в § 1.1. Однако теперь понимание этой важной функции значительно углубилось: мы выяснили, что периодограмма не только служит удобным инструментом при выявлении скрытых периодов, но и отражает всю частотную структуру процесса. Отметим, что функции $P_T(\nu)$ и $P_N(\nu)$ имеют одинаковую размерность, равную квадрату размерности изучаемого процесса $y(t)$ и отличаю-

шуюся от размерности спектральной плотности $g(\nu)$ множителем $[t]^{-1}$.

Соотношение (7) позволяет рассматривать непрерывный аналог периодограммы, точнее функцию $TP_T(\nu)$, в качестве оценки спектральной плотности процесса $g(\nu)$. Возникает естественный вопрос: нельзя ли и $P_N(\nu)$ использовать с той же целью? Как легко видеть, дискретность отсчетов не позволяет столь просто связать периодограмму и спектральную плотность, хотя $P_N(\nu)$, конечно, содержит информацию о спектре временного процесса. Дело в том, что для дискретных отсчетов появляется еще одна новая функция, описывающая распределение во времени моментов наблюдения. Учет этой функции сильно усложняет соотношение между $P_N(\nu)$ и $g(\nu)$; в частности, именно дискретностью наблюдений обусловлено явление подмены частот, отсутствующее при непрерывной регистрации. Обратимся к выводу соотношений, связывающих периодограмму $P_N(\nu)$ и спектр исходного непрерывного процесса. В данном случае процесс удобно представить в виде суммы детерминированного и случайного компонентов, как это сделано в (1.2.1). Поскольку спектры тренда и периодического процесса можно, вообще говоря, рассматривать единообразно, выбранное представление не уменьшает общности анализа, позволяя в то же время упростить выкладки. Более того, в качестве первого шага выясним влияние дискретности отсчетов отдельно на детерминированный и стохастический компоненты.

Детерминированный процесс. Предположим, что наблюдаемый процесс включает только центрированную неслучайную функцию $f(t)$, т.е. $y^{(0)}(t) = f(t)$. Полагая

$$D_N(\nu) \equiv \frac{1}{N} \sum_{k=0}^{N-1} e^{-i\cdot 2\pi\nu t_k} f(t_k), \qquad -\infty < \nu < \infty, \tag{1.6.13}$$

находим из (10) и (11):

$$P_N(\nu) = |D_N(\nu)|^2. \tag{1.6.14}$$

Функцию $D_N(\nu)$ называют *дискретным спектром* процесса $f(t)$. Если воспользоваться формулой замены аргумента (П.2.5) для δ-функции Дирака, то определение (13) можно представить в

виде

$$D_N(\nu) = \int_{-\infty}^{\infty} e^{-i\cdot 2\pi\nu t} f(t) \left[\frac{1}{N}\sum_{k=0}^{N-1} \delta(t-t_k)\right] dt, \qquad (1.6.15)$$

который ясно показывает, что дискретность процесса приводит к появлению под знаком интеграла дополнительного множителя

$$w_N(t) = \frac{1}{N}\sum_{k=0}^{N-1} \delta(t-t_k), \qquad (1.6.16)$$

называемого *временным окном данных*. При непрерывной регистрации в бесконечном промежутке вместо суммы δ-функций в (15) следует подставить 1, и мы приходим к обычному преобразованию Фурье процесса $f(t)$; непрерывная регистрация в течение конечного промежутка T предполагает замену указанной суммы функцией $r_T(t)$ - прямоугольным распределением (1.3.19). В случае же дискретной регистрации (15) можно формально трактовать как преобразование Фурье

$$D_N(\nu) = \int_{-\infty}^{\infty} e^{-i\cdot 2\pi\nu t} f_N(t)dt \qquad (1.6.17)$$

функции

$$f_N(t) = f(t)\, w_N(t). \qquad (1.6.18)$$

Умножение $f(t)$ на сумму δ-функций в (18) описывает процесс дискретизации отсчетов. Следует, однако, иметь в виду, что (18), как и все соотношения, в которые входит δ-функция, имеет операторный смысл и становится справедливым в обычном понимании только тогда, когда δ-функции находятся под знаком интеграла.

Согласно (1.3.9) и (1.3.10), преобразование Фурье произведения $f(t)\, w_N(t)$ равно свертке преобразований Фурье отдельно взятых функций $f(t)$ и $w_N(t)$, т.е. свертке $F(\nu)$ и функции

$$W_N(\nu) = \int_{-\infty}^{\infty} e^{-i\cdot 2\pi\nu t} w_N(t)dt = \frac{1}{N}\sum_{k=0}^{N-1} e^{-i\cdot 2\pi\nu t_k}, \qquad (1.6.19)$$

которую называют, следуя Блэкмену и Тьюки (1958), *спектральным окном данных*. Мы приходим к фундаментальному

соотношению

$$D_N(\nu) = \int_{-\infty}^{\infty} W_N(\nu-\nu')F(\nu')d\nu', \qquad -\infty < \nu < \infty, \tag{1.6.20}$$

связывающему между собой истинный спектр процесса $F(\nu)$, его дискретный спектр $D_N(\nu)$ и спектральное окно данных $W_N(\nu)$. Насколько нам известно, для общего случая неравномерных рядов соотношение (20) было найдено Фелжеттом (1972). Следствия (20), описывающие эффекты группировки данных, в особенно ясной форме обсуждались Дамингом (1975a, b), а затем и многими другими авторами.

Периодограмма детерминированного процесса $f(t)$, согласно (14) и (20), равна

$$P_N(\nu) = \left| \int_{-\infty}^{\infty} W_N(\nu-\nu')F(\nu')d\nu' \right|^2. \tag{1.6.21}$$

Мы видим, что дискретный характер отсчетов не позволяет связать $P_N(\nu)$ непосредственно со спектральной плотностью исходного процесса. Вместо этого следует сначала свернуть его спектр $F(\nu)$ со спектральным окном данных $W_N(\nu)$, и лишь затем вычислить квадрат модуля.

Перечислим некоторые свойства функций $D_N(\nu)$ и $W_N(\nu)$. Поскольку обе они - суть образы Фурье действительных функций, мы имеем:

$$D_N(-\nu) = D_N^*(\nu), \qquad W_N(-\nu) = W_N^*(\nu). \tag{1.6.22}$$

Далее, из (13) и (19) следует:

$$D_N(0) = \frac{1}{N}\sum_{k=0}^{N-1} f(t_k) = \bar{f}, \tag{1.6.23}$$

$$W_N(0) = 1, \quad |W_N(\nu)| \leqslant 1. \tag{1.6.24}$$

Если временной процесс центрирован, то $D_N(0) = 0$. Как легко видеть, $|W_N(\nu)|$ - симметричная функция частоты, а $\arg W_N(\nu)$ - несимметричная. Разумеется, периодограмма $P_N(\nu)$ и плотность спектрального окна данных $|W_N(\nu)|^2$ являются симметричными функциями частоты.

Мы увидим в гл. 2, что для равномерных временных рядов с шагом Δt максимальное значение $|W_N(\nu)| = 1$ достигается при $\nu = \pm l/\Delta t$, $l = 0, 1, 2, \ldots$, так что функция $|W_N(\nu)|$ при $N \gg 1$ напоминает гребенку с единичными зубцами, разделенными интервалом $1/\Delta t$ (см. рис. 2.1.1). Для неравномерных рядов зубцы при $\nu \neq 0$ меньше центрального, их расположение соответствует группировке данных во времени. Так, если моменты наблюдений модулированы некоторой частотой ν_w (обычно - частотой вращения Земли вокруг оси и/или частотой ее обращения вокруг Солнца), то $|W_N(\nu)|$ содержит максимумы на частотах $\pm\nu_w$, $\pm 2\nu_w, \ldots$ Это обстоятельство имеет прямое отношение к проблеме выявления скрытой периодичности.

В самом деле, предположим, что истинный спектр процесса содержит набор линий на частотах ν_1, $\nu_2, \ldots$, а при остальных значениях частоты $|F(\nu)|$ мал. Как показывает (20), при этом в первом приближении можно считать, что каждому максимуму истинного спектра отвечает слагаемое вида $F(\nu_n)W_N(\nu-\nu_n)$ в представлении дискретного спектра Фурье. Другими словами, спектральное окно данных как бы «цепляется» за каждый подъем истинного спектра. В итоге наблюдаемый спектр $D_N(\nu)$ и периодограмма $P_N(\nu)$ искажены большим количеством ложных пиков, связанных с группировкой моментов наблюдения (см. пример 1.6.2 ниже). Мы вновь пришли к явлению подмены частот, но уже для общего случая неравномерных временных рядов.

Из соотношения (21) нетрудно усмотреть основную интерпретацию функции $|W_N(\nu)|^2$. Именно, плотность спектрального окна данных есть периодограмма функции $f(t) \equiv 1$, которая наблюдается в моменты времени t_0, t_1, ..., t_{N-1}.

Каким же образом на основании (20) можно было бы восстановить истинный спектр процесса $F(\nu)$? Позитивного ответа добиться нелегко, однако сразу можно указать путь, представляющийся бесперспективным. Речь идет о прямом решении интегрального уравнения (20) стандартными методами интегральных преобразований для уравнений типа свертки. Во-первых, (20) относится к классу некорректно поставленных задач, и небольшая добавка шума (а он обязательно присутству-

ет при измерениях) существенно влияет на характер получаемого решения (см. § 3.8). Во-вторых, функция $w_N(t)$ - обратное преобразование Фурье ядра интегрального уравнения (20) - равна нулю для тех моментов времени, когда не было наблюдений. Поэтому даже формальное обращение (20) сталкивается с трудностями - «... ничего нельзя сделать, чтобы смягчить потерю наблюдательной информации» (Фелжетт, 1972).

Конструктивный подход к проблеме получения информации о спектре исходного процесса изложен в гл.2 и 3. Для приближенной оценки $|F(\nu)|^2$ часто используется прием, введенный в астрономическую практику, по-видимому, Фитчем (1967), и получивший широкое распространение после детального обсуждения в статье Диминга (1975а). Именно, на основании данных наблюдений можно построить периодограмму $P_N(\nu)$ и отдельно - плотность спектрального окна данных $|W_N(\nu)|^2$, а затем судить о реальности деталей периодограммы на основании соображений типа высказанных выше о том, что спектральное окно «цепляется» за максимумы истинного спектра. Фактически это означает интуитивное решение некорректной задачи (20), (21).

Пример 1.6.1. Пусть $f(t)$ есть гармоническое колебание (1.4.2). Его истинный спектр $F(\nu)$ определяется соотношением (1.4.14), при подстановке которого в (20) получается следующее выражение для дискретного спектра гармонического процесса:

$$D_N(\nu) = \frac{A}{2}\,[e^{-i\varphi_0} W_N(\nu-\nu_0) + e^{i\varphi_0} W_N(\nu+\nu_0)], \qquad (1.6.25)$$

справедливое в общем случае произвольного неравномерного временного ряда. Соответствующая (25) периодограмма имеет вид

$$P_N(\nu) = \frac{A^2}{4}\,[|W_N(\nu-\nu_0)|^2 + |W_N(\nu+\nu_0)|^2 + 2|W_N(\nu-\nu_0)||W_N(\nu+\nu_0)|\cos\gamma_N(\nu)], \qquad (1.6.26)$$

где

$$\gamma_N(\nu) = \arg W_N(\nu-\nu_0) - \arg W_N(\nu+\nu_0) - 2\varphi_0. \qquad (1.6.27)$$

Заметим, что мы отвлеклись пока от точного центрирования ряда, играющего роль лишь при весьма небольшой его длине.

Пример 1.6.2. На рис. 1.6.1 представлены плотность спектрального окна и периодограмма переменной звезды *CF Cas*, вычисленные нами на основании данных Бердникова (1986). Для того, чтобы не вводить годичную модуляцию, мы ограничились рядом наблюдений между моментами *JD Hel* (2445) 509 - 707. Этот временной отрезок продолжительностью T~200 дней включает лишь 28 отсчетов (табл. П. 7.5), т.е. в среднем интервал времени между измерениями составил около семи дней. Как и следовало ожидать, $|W_N(\nu)|^2$ содержит последовательно убывающие максимумы на частотах, кратных суткам. Основному периоду звезды $P_0 \simeq 4,9$ сут. соответствует максимум периодограммы на частоте $\nu_0 \simeq 0,204$ цикл/сут. Однако присутствуют и пики, отвечающие сопряженным частотам (1.1.17). Как уже говорилось выше, в первом приближении влияние сопряженных частот на периодограмму может быть изучено последовательным наложением графика $|W_N(\nu)|^2$ и симметричного ему графика по отрицательной области частот на основные гармоники $\pm \nu_0$ исходного спектра.

Случайный процесс. Рассмотрим теперь периодограмму для другого крайнего случая, когда наблюдаемый прицесс включает только стохастический компонент. Полагая $y^{(0)}(t) = \xi(t)$, находим из (10) и (11):

$$P_N(\nu) = \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} e^{-i \cdot 2\pi\nu(t_k - t_l)} \xi(t_k)\xi(t_l). \qquad (1.6.28)$$

Разумеется, периодограмма случайного процесса сама является случайной функцией. Ее среднее значение равно

$$\langle P_N(\nu) \rangle = \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} e^{-i \cdot 2\pi\nu(t_k - t_l)} \psi(t_k - t_l), \qquad (1.6.29)$$

где мы воспользовались определением (1.5.4). Подставим далее вместо $\psi(\tau)$ ее выражение (1.5.13) через $g(\nu)$. Тогда

$$\langle P_N(\nu) \rangle = \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} e^{-i \cdot 2\pi\nu(t_k - t_l)} \int_{-\infty}^{\infty} e^{i \cdot 2\pi\nu'(t_k - t_l)} g(\nu')\,d\nu'. \qquad (1.6.30)$$

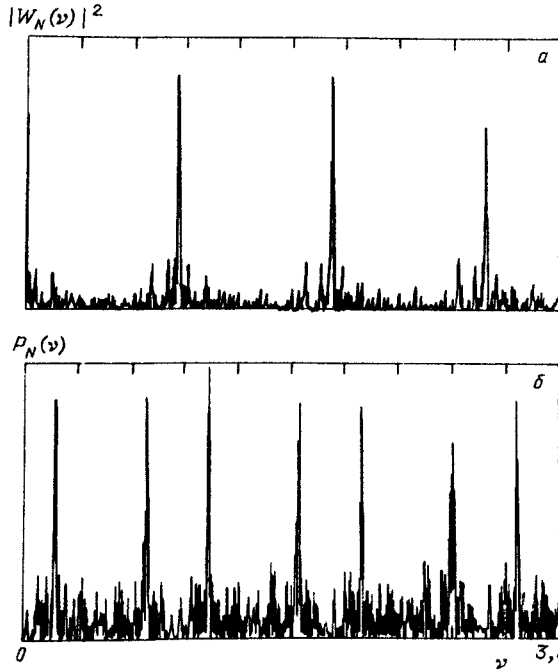


Рис. 1.6.1. Плотность спектрального окна данных (а) и периодограмма (б) для оценок блеска звезды CF Cas

Принимая во внимание определение (19) спектрального окна данных, получаем окончательно:

$$\langle P_N(\nu)\rangle = \int_{-\infty}^{\infty} |W_N(\nu-\nu')|^2 g(\nu')d\nu'. \qquad (1.6.31)$$

Это выражение следует сравнить с аналогичным выражением (21) для детерминированных процессов. Существенная разница между (31) и (20),(21) заключается в том, что теперь среднее значение периодограммы связано непосредственным образом со спектральной плотностью $g(\nu)$ уравнением типа свертки. Для детерминированных процессов такая связь имеется между дискретным и непрерывным спектрами. Что касается решения интегрального уравнения (31), то здесь справедливы те же замечания, которые были высказаны относительно решения (20) и (21). Заметим, что приведенная выше основная интерпретация $|W_N(\nu)|^2$ как периодограммы постоянной функции сохраня-

ется и в рамках спектральной теории случайных процессов.

Сумма детерминированного и случайного процессов. Изучение указанных процессов по отдельности, несомненно, оправдано методически. Однако в действительности мы всегда имеем дело либо с некоторой комбинацией детерминированного и случайного процессов, либо только с последним из них. Рассмотрим поэтому наблюдаемый ряд $\{y_k^{(o)}\}$ в виде суммы центрированных рядов $\{f_k\}$, $\{\xi_k\}$ и найдем выражение периодограммы (10) этого суммарного ряда через спектральные характеристики компонентов. Мы не будем приводить здесь соответствующие выкладки, поскольку они аналогичны описанным выше. Искомое соотношение между средней периодограммой $\langle P_N(\nu)\rangle$, спектром детерминированного процесса $F(\nu)$ и спектральной плотностью стохастического компонента $g(\nu)$ имеет вид

$$\langle P_N(\nu)\rangle = \left|\int_{-\infty}^{\infty} W_N(\nu-\nu')F(\nu')d\nu'\right|^2 + \\ + \int_{-\infty}^{\infty} \left|W_N(\nu-\nu')\right|^2 g(\nu')d\nu', \qquad (1.6.32)$$

т.е. средняя периодограмма (но не ее конкретная реализация) представляет собой сумму периодограммы (21) детерминированного процесса и средней периодограммы (31) стохастического процесса. Все высказанные соображения о решении (20) сохраняют силу и относительно решения (32).

Спектральное разрешение Обратим внимание на то обстоятельство, что дискретный характер наблюдений мало влияет на спектральное разрешение $\delta\nu$, которое составляет для процесса длительности T величину порядка T^{-1}. Конкретное распределение моментов наблюдения $\{t_k\}$ по интервалу T оказывает воздействие на форму спектрального отклика, однако общая ширина линии остается порядка T^{-1}. Для суждения о форме спектральных линий и для более точного определения положения максимумов шаг по частоте при вычислении периодограммы следует принять равным

$$\delta\nu = \frac{1}{KT}, \qquad (1.6.33)$$

где параметр $K \simeq 3 \div 10$ (в частности, периодограммы § 1.1 рассчитаны при $K = 10$). Постоянный шаг (33) в области частот соответствует переменному шагу

$$\delta P = \frac{P^2}{KT} \qquad (1.6.34)$$

в области периодов, что и предопределяет удобства, связанные с использованием шкалы частот.

Выражение (26) показывает, что положение максимума периодограммы не вполне точно совпадает с истинной частотой гармонического колебания. Еще более значительное смещение частоты наблюдается при наличии двух колебаний с близкими частотами (Лумос и Диминг, 1978). Детальное исследование относящихся сюда вопросов выполнил Ковач (1981).

§ 1.7. Два подхода к проблеме поиска периодичности

При изучении ряда A в § 1.1 было найдено приближенное значение $\hat{P}_0$ истинного периода P_0, а затем построена фазовая диаграмма, т.е. последовательные отрезки временного ряда $\{y_k\}$ длиной $\hat{P}_0$ были наложены друг на друга и рассмотрена зависимость y_k от фазы (1.1.18). Результат операции представлен на рис. 1.1.3. Цель построения фазовой диаграммы заключалась в данном случае в том, чтобы приближенно определить форму истинной периодической зависимости $y(x)$.

Представляется вполне очевидным, что при $\hat{P}_0$, близком к истинному периоду P_0, точки на фазовой диаграмме располагаются наиболее упорядоченным образом. В противном случае разброс точек будет обусловлен не только неизбежным шумом, но и причинами систематического характера, т.е. будет превышать разброс, соответствующий сворачиванию ряда с периодом P_0. Таким образом, разброс точек на фазовой диаграмме может служить показателем близости какого-либо значения периода к истинному периоду P_0. Но тогда можно вообще отказаться от нахождения приближенного значения $\hat{P}_0$ с помощью периодограммы и предложить следующую процедуру выделения скрытой периодической зависимости:

1) определить некоторую числовую характеристику Γ степени упорядоченности расположения точек на фазовой диаграмме;

2) построить фазовые диаграммы изучаемого временного ряда для всех интересующих нас значений пробного периода P (или частоты $\nu = 1/P$), и для каждой фазовой диаграммы вычислить соответствующую характеристику $\Gamma(\nu)$;

3) выбрать то значение Γ_0, которое отвечает наиболее упорядоченному расположению точек (обычно это минимальное или максимальное - в зависимости от определения - значение Γ). Соответствующая Γ_0 частота ν_0 и укажет на искомое значение этого параметра, а расстояние $\Gamma(\nu_0) \equiv \Gamma_0$ от типичных для других частот значений $\Gamma(\nu)$ - на степень надежности найденной периодической зависимости.

В качественном виде описанный подход настолько давно употребляется в астрономии, что первоисточник трудно указать. До середины 60-х годов фазовые диаграммы изменений блеска или лучевых скоростей звезд строились вручную, а степень упорядоченности точек оценивалась на глаз (что имеет свои преимущества). Разумеется, при этом удавалось перебрать лишь небольшое число пробных периодов, и успех работы определялся опытом исследователя или удачей. Наглядно перебор пробных периодов можно представить следующим образом. Предположим, что отсчеты временного ряда $\{y(t_k)\}$ нанесены на прозрачную бумагу, которая затем свернута в цилиндрический рулон в направлении оси времени. Осматривая рулон со всех сторон (для того чтобы избежать сквозного просмотра, в середину рулона следует поместить лист бумаги), мы получим возможность судить о фазовой диаграмме ряда для пробного периода $P = \pi D$, где D - диаметр рулона. Меняя D, можно добиться наиболее тесного расположения точек около некоторой средней «кривой блеска» и определить тем самым искомый период. С появлением электронных компьютеров представилась возможность существенно увеличить объем расчетов и как следствие - повысить надежность выводов. Однако для этого необходимо было формализовать описанную процедуру и - прежде всего - дать надлежащее определение параметра $\Gamma(\nu)$.

Первый шаг в этом направлении был сделан Лафлером и Кинманом (1965). Они предложили принять в качестве характеристики разброса точек на фазовой диаграмме сумму квадратов разностей ординат последовательных точек на ней, т.е.

известный параметр Аббе (см. Большев и Смирнов, 1965). Подробное обсуждение статистики Аббе-Лафлера-Кинмана $\Theta_N(\nu)$ дано в § 4.1, здесь нам достаточно указать, что при близком в среднем расположении последовательных точек на фазовой диаграмме $\Theta_N \ll 1$, в то время как при независимом случайном их размещении ожидаемое значение Θ_N близко к 1. Обе эти характерные ситуации иллюстрируются приводимым ниже примером.

Пример 1.7.1. Обратимся вновь к наблюдениям переменной звезды CF Cas, выполненным Бердниковым (1986). В примере 1.6.2 эти данные были рассмотрены с точки зрения спектральной теории, найдем теперь зависимость статистики $\Theta_N(\nu)$ в интервале частот $0 \le \nu \le 3{,}5$ цикл/сут с шагом $\delta\nu = 0{,}001$ цикл/сут (рис. 1.7.1). Как и ранее, основной частоте $\nu_0 \simeq 0{,}204$ цикл/сут отвечает наиболее выделяющийся экстремум, в данном случае - минимум функции $\Theta_N(\nu)$. Менее глубокие минимумы имеются и на сопряженных частотах $\nu_l = \pm\nu_0 \pm l/\Delta t$, $l = 1, 2, \ldots$ Следует, однако, обратить внимание на существование минимумов $\Theta_N(\nu)$ на частотах субгармоник $\nu_l^{(m)} = \nu_l/m$, $m = 1, 2, \ldots$, соответствующих периодам, кратным сопряженным. Наша прежняя статистика - периодограмма Шустера $P_N(\nu)$ - не содержит экстремумов на частотах субгармоник.

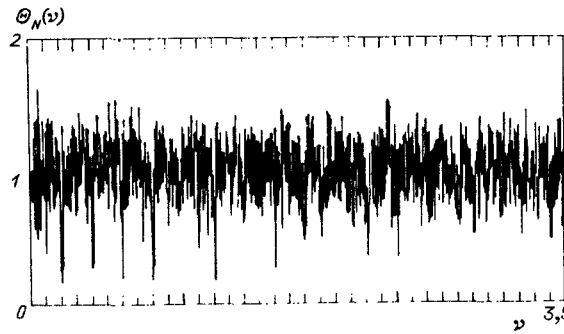


Рис. 1.7.1. Периодограмма Аббе-Лафлера-Кинмана для оценок блеска CF Cas

На рис. 1.7.2 представлены фазовые диаграммы измерений блеска CF Cas (28 отсчетов) для частоты основного колебания ν_0 и некоторой произвольно взятой частоты.

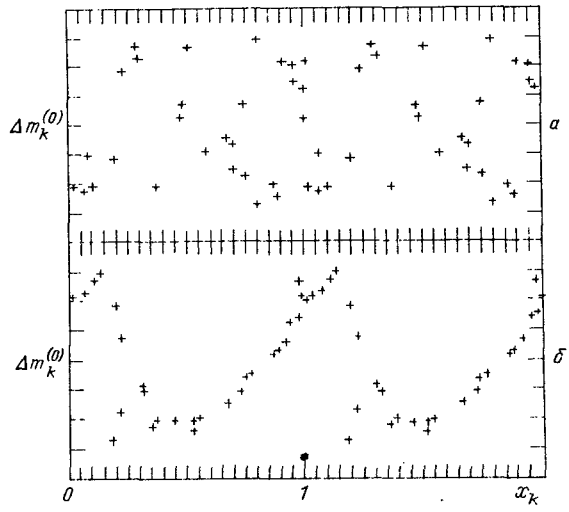


Рис. 1.7.2. Фазовые диаграммы оценок блеска CF Cas для ν = = 0,300 цикл/сут (а) и ν = 0,204 цикл/сут (б)

Хотя конкретный выбор меры рассеяния точек на фазовой диаграмме, сделанный Лафлером и Кинманом, оказался удачным, он не остался единственным. За два последних десятилетия было предложено несколько обобщений статистики $\Theta_N(\nu)$, учитывающих различные свойства распределения точек на фазовой диаграмме. Некоторые из этих обобщений рассматриваются в гл. 4. Все они, как и метод Аббе-Лафлера-Кинмана, являются разновидностями так называемого *метода сворачивания эпох* (epoch folding). Нередко употребляется и название *метод наложения эпох.*

Возникает естественный вопрос: как соотносятся между собой классический метод спектрального анализа и метод сворачивания эпох?

Прежде всего обратим внимание на то очевидное обстоятельство, что периодограмма также основана на предварительном сворачивании ряда с набором пробных частот. Действительно, вернемся к вспомогательным функциям $U_N(\nu)$ и $V_N(\nu)$, определенным соотношениями (1.1.11). Представим νt_k в виде

$$\nu t_k = \mathrm{INT}(\nu t_k) + x_k(\nu) \ , \qquad (1.7.1)$$

где INT(z) - целая часть числа z, а $x_k(\nu) \equiv \mathrm{FR}(\nu t_k)$ - дробная часть νt_k, т.е. фаза, отвечающая моменту наблюдений t_k. Ввиду периодичности тригонометрических функций (1.1.11) можно переписать следующим образом:

$$U_N(\nu) = \frac{1}{N}\sum_{k=0}^{N-1}\cos[2\pi x_k(\nu)]y_k^{(0)},$$
$$V_N(\nu) = \frac{1}{N}\sum_{k=0}^{N-1}\sin[2\pi x_k(\nu)]y_k^{(0)}. \qquad (1.7.2)$$

Мы видим, что $U_N(\nu)$ с точностью до множителя N^{-1} есть коэффициент корреляции исследуемого ряда с функцией $\cos(2\pi x)$, заданной на интервале фаз $0 \leqslant x < 1$, а $V_N(\nu)$ - аналогичный коэффициент корреляции с функцией $\sin(2\pi x)$. Периодограмма же (1.1.13) учитывает корреляцию ряда с гармоническим колебанием $\cos(2\pi x - \varphi_0)$ при оптимальном выборе начальной фазы φ_0 согласно (1.1.12).

Итак, периодограмма также характеризует упорядоченность распределения точек на фазовой диаграмме, но форма искомого упорядоченного расположения задана наперед - она описывается тригонометрической функцией. Методы, не связанные таким ограничением (например, метод Аббе-Лафлера-Кинмана), называются *непараметрическими* (гл. 4). Ввиду сказанного все статистики, характеризующие фазовую диаграмму временного ряда при различных значениях пробного периода, т.е. $P_N(\nu)$, $\Theta_N(\nu)$ и т.п. мы будем называть единообразно - периодограммами, различая их по собственным именам.

К достоинствам метода сворачивания эпох следует отнести его чувствительность при обнаружении колебаний, форма которых сильно отличается от синусоидальной. В этом случае спектральная плотность содержит высокие гармоники (обертоны), выделение которых из шумов нередко затруднительно. Именно так и произошло для ряда А в примере § 1.1. Однако отмеченная особенность метода сворачивания эпох имеет и негативную сторону: он выделяет вместе с основной и тривиальные периодичности с периодами, кратными основному, т.е.

частоты $\nu_0/2$, $\nu_0/3$,... Мы уже столкнулись с этим при обсуждении примера 1.7.1.

Подобно тому как существует множество вариантов метода сворачивания эпох, отличающихся лишь конкретным видом статистики $\Gamma(\nu)$, можно предложить другие, помимо периодограммы Шустера, статистики, сравнивающие изучаемый временной ряд с некоторым заданным наперед колебанием. Например, если изучается переменность блеска звезды, относящейся к группе объектов с весьма сходными кривыми блеска $f(x; \alpha_1,\ldots,\alpha_m)$, то в качестве $\Gamma(\nu)$ можно выбрать сумму квадратов уклонений отсчетов временного ряда от функции $f(x; \alpha_1,\ldots,\alpha_m)$, оставив неизвестными лишь конечное число параметров $\alpha_1,\ldots,\alpha_m$. Эффективность периодограммы Шустера обусловлена тем важным фактором, что тригонометрические функции $\sin(2\pi n\nu_o x)$, $\cos(2\pi n\nu_o x)$, $n = 1, 2,\ldots$ образуют *полную систему*, по которой можно разложить практически любую функцию.

Вместе с тем существует много полных систем функций, так что не менее полезными, чем периодограмма Шустера, а в отдельных ситуациях и более эффективными могут оказаться другие статистики. Особого внимания, как нам представляется, заслуживает разложение по функциям Хаара (эти функции описаны, например, Соболем, 1969). Периодограмма Хаара предпочтительнее других и с вычислительной точки зрения.

ГЛАВА 2

РАВНОМЕРНЫЕ ВРЕМЕННЫЕ РЯДЫ

В предыдущей главе было показано, что спектральный анализ временных рядов опирается на изучение периодограммы $P_N(\nu)$ и спектрального окна данных $W_N(\nu)$. Эти функции, заданные соответственно формулами (1.6.12) и (1.6.19), можно найти непосредственно по результатам наблюдений; особое значение $W_N(\nu)$ определяется интегральным уравнением (1.6.32), связывающим среднюю периодограмму со спектром $F(\nu)$ детерминированного компонента и спектральной плотностью $g(\nu)$ стохастического компонента процесса.

Настоящая глава посвящена более подробному рассмотрению частного случая, когда промежутки времени между моментами наблюдений равны между собой, т.е. *равномерным временным рядам*. Поскольку продолжительность одного цикла наземных наблюдений не превышает нескольких часов, такие ряды встречаются в астрономической практике значительно реже, чем в прикладных исследованиях. Тем не менее равномерные временные ряды заслуживают внимательного изучения. Во-первых, они естественным образом выделены из множества неравномерных рядов, каждый из которых, можно сказать, неравномерен по-своему. Затем, сравнительная простота равномерного ряда позволяет продвинуться дальше в его изучении и получить методическую основу для исследования более общих случаев. Наконец, равномерным рядам присущи некоторые особенности, которые необходимо знать при интерпретации конкретных данных наблюдений.

Как уже отмечалось в предисловии, имеющаяся монографическая литература рассматривает почти исключительно равно-

мерные временные ряды (см., в частности, монографию Серебренникова и Первозванского (1965), в которой описаны методы, применявшиеся начиная с конца 18-го века, и современные руководства Кендалла и Стюарта (1966), Т. Андерсона (1971), и Бриллинджера (1975)). Это обстоятельство значительно облегчает нашу задачу, позволяя в данной главе ограничиться наиболее характерными чертами теории и сосредоточить внимание на вопросах, приобретающих значение при дальнейшем изучении неравномерных рядов.

Повсюду в гл. 2 предполагается для простоты, что среднее значение процесса точно известно, а временные ряды соответствующим образом центрированы. При небольшой длине ряда следует учитывать отличие выборочного и теоретических средних; связанные с этим вопросы подробно рассматриваются Т. Андерсоном (1971).

§ 2.1. Основные соотношения

Спектральное окно. Начнем изучение равномерных рядов с выяснения свойств функции $W_N(\nu)$ - спектрального окна данных. Эта функция не зависит от значений, принимаемых рядом, и характеризует только распределение моментов наблюдения во времени. Полагая $t_k = k\Delta t$, $k = 0, 1, \ldots, N-1$ в (1.6.19), находим:

$$W_N(\nu) = \frac{1}{N}\sum_{k=0}^{N-1} e^{-i\cdot 2\pi\nu\Delta t k} = \frac{1}{N}\,\frac{1-e^{-i\cdot 2\pi x N}}{1-e^{-i\cdot 2\pi x}}, \qquad (2.1.1)$$

где $x = \nu\,\Delta t$ - безразмерная частота. Отсюда получаем для модуля комплексной функции $W_N(\nu)$ выражение

$$|W_N(\nu)| = \left|\frac{\sin(\pi N x)}{N\sin(\pi x)}\right|, \qquad (2.1.2)$$

и для ее аргумента

$$\arg W_N(\nu) = \begin{cases} -\pi(N-1)x, & \text{если } x \text{ или } Nx \text{ не равны} \\ & \text{целому числу,} \\ 0 & \text{в противном случае.} \end{cases} \qquad (2.1.3)$$

Функция

$$\Phi_N(x) = \frac{\sin(\pi N x)}{N\sin(\pi x)}, \quad -\infty < x < \infty \qquad (2.1.4)$$

хорошо известна из теории дифракции электромагнитных волн. Основные ее свойства перечислены в Приложении 3, графики $\phi_N^2(x)$ при различных значениях N представлены на рис. 2.1.1.

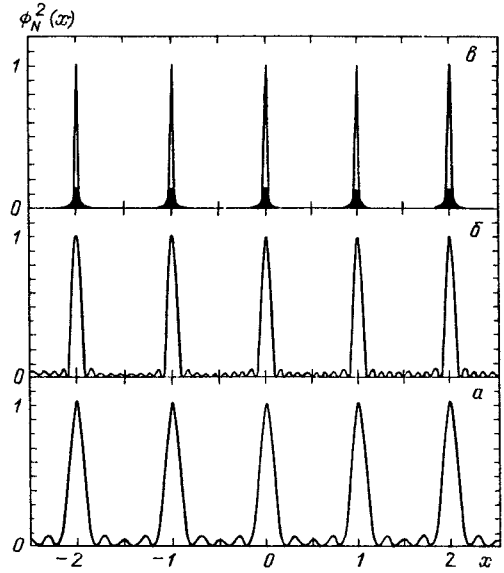


Рис. 2.1.1. Плотность спектрального окна данных равномерного временного ряда в зависимости от безразмерной частоты x при N = 5 (а), 10 (б) и 50 (в)

Мы видим, что при $N \gg 1$ плотность спектрального окна имеет вид единичной «гребенки» с шириной зубцов $\delta\nu \simeq (N\Delta t)^{-1}$ и расстоянием между ними $\nu_s = (\Delta t)^{-1}$, равным частоте отсчетов выборки. Такого вида $|W_N(\nu)|^2$ и следовало ожидать на основании обсуждавшейся в § 1.6 интерпретации плотности спектрального окна как периодограммы функции, тождественно равной 1. Действительно, функция $f(t) \equiv 1$ точно удовлетворяет совокупности N равноотстоящих единичных отсчетов, расположенных через интервал Δt, и потому $|W_N(0)|^2 = 1$. Однако через эти же отсчеты можно точно провести и синусоиды с частотами $\nu_l = \pm l/\Delta t$, $l = 1, 2, \ldots$, так что $|W_N(\nu_l)|^2 = 1$ для всех целых значений l. Ширина каждого максимума составляет величину порядка T^{-1}, как это мы уже обсуждали в § 1.3 и § 1.4.

Дискретный спектр. Рассмотрим дискретный спектр $D_N(\nu)$

детерминированного процесса $f(t)$. Полагая $t_k = k\Delta t$ в (1.6.13), находим:

$$D_N(\nu) = \frac{1}{N}\sum_{k=0}^{N-1} e^{-i\cdot 2\pi\nu\Delta t k}\, f(t_k), \quad -\infty < \nu < \infty . \tag{2.1.5}$$

Некоторые общие свойства $D_N(\nu)$ уже обсуждались в § 1.6, при равномерном распределении моментов наблюдения эта функция характеризуется дополнительной симметрией.

Заметим прежде всего, что $D_N(\nu)$ – периодическая функция с периодом Δt^{-1}:

$$D_N(\nu \pm l/\Delta t) = D_N(\nu), \quad l = 0, 1, \ldots \tag{2.1.6}$$

Это позволяет ограничиться при расчетах частотным интервалом $0 \leqslant \nu \leqslant 1/\Delta t$. В этой области $D_N(\nu) = D_N^*(1/\Delta t - \nu)$, так что периодограмма $P_N(\nu) = |D_N(\nu)|^2$ симметрична относительно частоты Найквиста $\nu_c = (2\Delta t)^{-1}$:

$$P_N(1/\Delta t - \nu) = P_N(\nu), \quad 0 \leqslant \nu \leqslant 1/\Delta t . \tag{2.1.7}$$

Пример 2.1.1. Периодограмма гармонического колебания весьма просто выражается через спектральное окно данных формулами (1.6.26), (1.6.27). В случае равномерного ряда мы имеем на основании (2)-(4):

$$P_N(\nu) = \frac{A^2}{4}\left\{\phi_N^2(x-x_0) + \phi_N^2(x+x_0) + 2\left|\phi_N(x-x_0)\phi_N(x+x_0)\right|\cos[2\pi(N-1)x_0 - 2\varphi_0]\right\}, \tag{2.1.8}$$

где $x = \nu\Delta t$, $x_0 = \nu_0\Delta t$; A, ν_0 и φ_0 – соответственно амплитуда, частота и начальная фаза колебания. При $N \gg 1$ в интервале $0 \leqslant x \leqslant 0{,}5$ фактически заметна лишь одна спектральная линия вблизи основной частоты:

$$P_N(\nu) \simeq \frac{A^2}{4}\phi_N^2(x-x_0) , \tag{2.1.8'}$$

форма которой иллюстрируется рис 2.1.1.

Выше неоднократно указывалось, что спектральное разрешение составляет величину порядка $\delta\nu = T^{-1} = [(N-1)\Delta t]^{-1}$. Поэтому $D_N(\nu)$ можно рассматривать на дискретной сетке значений частоты

$$\nu_m = \frac{m}{(N-1)\Delta t}, \quad m = 0, 1, \ldots, N-1, \tag{2.1.9}$$

где

$$D_N(\nu_m) = \frac{1}{N} \sum_{k=0}^{N-1} e^{-i \cdot \frac{2\pi}{N-1} mk} f(t_k). \tag{2.1.10}$$

Полагая для краткости

$$f(t_k) \equiv f_k, \quad D_N(\nu_m) \equiv D_m, \tag{2.1.11}$$

находим окончательно

$$D_m = \frac{1}{N} \sum_{k=0}^{N-1} f_k e^{-i \cdot \frac{2\pi}{N-1} mk}, \quad m = 0, 1, \ldots, N-1, \tag{2.1.12}$$

В этой форме дискретное преобразование Фурье (ДПФ) обычно и применяется при изучении равномерных временных рядов. Подобно обычному преобразованию Фурье формула (12) допускает обращение:

$$f_k = \sum_{m=0}^{N-1} D_m e^{i \cdot \frac{2\pi}{N-1} km}, \quad k = 0, 1, \ldots, N-1. \tag{2.1.13}$$

Для доказательства (13) достаточно принять во внимание равенство

$$\frac{1}{N} \sum_{m=0}^{N-1} e^{i \cdot \frac{2\pi}{N-1} rm} \cdot e^{-i \cdot \frac{2\pi}{N-1} mk} = \delta_{nk}, \tag{2.1.14}$$

где δ_{nk} - символ Кронекера.

Если распространить значения m на все целые числа, то условие периодичности (6) принимает вид

$$D_{m \pm l(N-1)} = D_m, \quad l = 0, 1, \ldots; \quad m = 0, \pm 1, \pm 2, \ldots \tag{2.1.15}$$

Мы можем также рассматривать совокупность $\{f_k\}$, формально определенную соотношением (13), для всех целых значений k. Как легко видеть, при этом выполняется равенство

$$f_{k \pm l(N-1)} = f_k, \quad l = 0, 1, \ldots; \quad k = 0, \pm 1, \pm 2, \ldots \tag{2.1.16}$$

Отметим еще полезные соотношения

$$D_{-m} = D_m^*, \quad D_m = D_{N-1-m}^*, \quad m = 0, \pm 1, \pm 2, \ldots, \tag{2.1.17}$$

вытекающие из (7) и (1.6.22).

Таким образом, формулы (12) и (13) допускают периодическое продолжение на все множество целых чисел. Это обстоятельство позволяет применить для вычисления совокупности $\{D_m\}$ алгоритм *быстрого преобразования Фурье*, значительно ускоряющий расчеты. Поскольку указанный алгоритм в настоящее время широко распространен, мы не будем его излагать. Отметим лишь, что эффективность быстрого преобразования Фурье особенно велика при значительной длине временного ряда: коэффициент ускорения по сравнению с прямым вычислением по формуле (12) составляет $N/\log_2 N$ (см., например, Кули и др., 1986).

Выполнив ДПФ данных наблюдений, можно найти и дискретные отсчеты периодограммы:

$$P_m \equiv P_N(\nu_m) = |D_m|^2, \quad m = 0, 1, \ldots, N-1, \qquad (2.1.18)$$

где в действительности ввиду симметрии P_m относительно частоты Найквиста достаточно принять максимальное значение m равным $N/2$ при четном N и равным $(N+1)/2$ - при нечетном N. В дискретном случае справедливо соотношение

$$\sum_{m=0}^{N-1} P_m = \frac{1}{N} \sum_{k=0}^{N-1} f_k^2 \qquad (2.1.19)$$

(*теорема Парсеваля*), аналогичное равенствам (1.3.15), (1.4.30) и (1.5.14) для детерминированных, периодических и случайных непрерывных функций соответственно (теорема Рэлея).

Истинный спектр. На основании данных наблюдений мы можем вычислить значения функций $D_N(\nu)$ и $W_N(\nu)$. Истинный спектр изучаемого процесса $F(\nu)$ связан с этими функциями интегральным уравнением (1.6.20), и мы должны исследовать возможность получения информации относительно $F(\nu)$ в случае равномерного ряда. Для понимания сути дела полезно рассмотреть (1.6.20) при $N \gg 1$, что, впрочем, представляет наибольший интерес и с точки зрения практики. Как показывает выполненный в начале данного параграфа анализ, при $N \gg 1$ плотность спектрального окна можно считать равной нулю между «зубцами», каждый из которых аппроксимируется прямоугольником единичной высоты с шириной порядка $(N\Delta t)^{-1}$. Но

тогда уравнение (1.6.20) принимает вид

$$D_N(\nu) \simeq \sum_{l=-\infty}^{\infty} F(\nu - l/\Delta t)\frac{1}{N\Delta t}\,, \quad N \gg 1\,. \qquad (2.1.20)$$

Итак, дискретный спектр равномерного временного ряда $\{f_k\}$ представляет собой наложение бесконечного числа спектров $F(\nu)$ исходного непрерывного процесса $f(t)$, последовательно сдвинутых друг относительно друга на одно и то же расстояние Δt^{-1}.

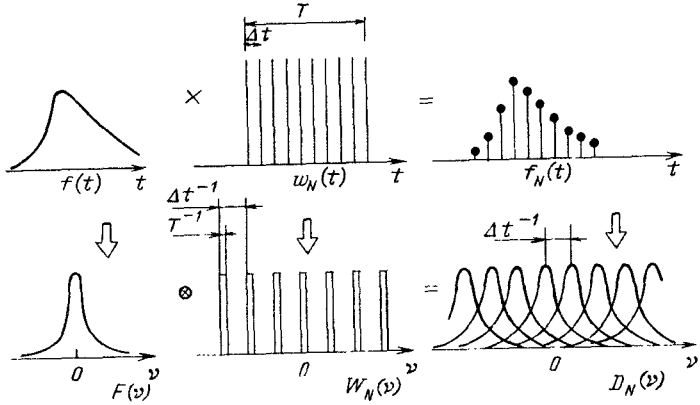


Рис. 2.1.2. Дискретизация непрерывного процесса $f(t)$ при равномерном временном окне данных $w_N(t)$ и соответствующая картина в частотной области, иллюстрирующая наложение спектров. Символы ⊗ и ⇒ означают свертку функций и преобразование Фурье соответственно

Наглядно процесс дискретизации непрерывной функции иллюстрируется рис. 2.1.2, который аналогичен соответствующему рисунку из монографии Д. Грея (1976).

Если $|F(\nu)|$ убывает достаточно быстро, а именно, если можно считать

$$F(\nu) = 0 \quad \text{при} \quad |\nu| > \nu_c \equiv (2\Delta t)^{-1}, \qquad (2.1.21)$$

то в сумме (20) отдельные спектры не перекрываются (см. рис. 2.1.2). Поэтому можно выбрать один из спектров (напри-

мер, тот, который находится в области частот, меньших граничной частоты Найквиста ν_c) и, обратив преобразование Фурье, восстановить временной процесс $f(t)$. Невыполнение условия (21) приводит к перекрытию отдельных спектров в сумме (20) тем более значительному, чем шире спектр исходного процесса $F(\nu)$ по сравнению с областью $|\nu| \leqslant \nu_c$. При этом знания $D_N(\nu)$, вообще говоря, недостаточно для однозначного восстановления исходного спектра $F(\nu)$, а следовательно, и временного процесса $f(t)$.

По сути дела перекрытие спектров в (20) совпадает с явлением наложения частот (aliasing), которое обсуждалось в § 1.1. Действительно, рассмотрим некоторую частоту $\nu < \nu_c$ (рис. 2.1.2.). Вклад в $D_N(\nu)$ от основного, несмещенного спектра процесса есть $F(\nu)$; каждый из спектров, расположенных правее, дает вклад вида $F(\nu - l/\Delta t)$, $l = 1, 2, \ldots$, а каждый из спектров, расположенных левее, - вида $F(\nu + l/\Delta t)$, $l = 1, 2, \ldots$ Таким образом, на частоте ν суммируются все детали исходного спектра $F(\nu)$ с частотами $\nu_l = \nu \pm l/\Delta t$, $l = 0, 1, \ldots$, что совпадает с последовательностью сопряженных частот (1.1.17).

Выборочная автоковариационная функция и периодограмма. Вернемся к случаю произвольного временного ряда, который может включать в себя как детерминированный, так и стохастический компоненты. При равномерном распределении моментов наблюдения выражения (1.6.12) для периодограммы принимают вид:

$$P_N(\nu) = U_N^2(\nu) + V_N^2(\nu),$$

$$U_N(\nu) = \frac{1}{N} \sum_{k=0}^{N-1} \cos(2\pi x k) y_k^{(0)}, \qquad (2.1.22)$$

$$V_N(\nu) = \frac{1}{N} \sum_{k=0}^{N-1} \sin(2\pi x k) y_k^{(0)},$$

где $x = \nu \Delta t$. Приведем другую форму периодограммы, привлекавшую особое внимание с точки зрения ускорения вычислений до широкого распространения быстрого преобразования Фурье, а теперь имеющую, главным образом, теоретическое значение.

Произведем с этой целью в (22) фактическое возведение

сумм $U_N(\nu)$ и $V_N(\nu)$ во вторую степень:

$$P_N(\nu) = \frac{1}{N^2}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1}\cos\,[2\pi(k-l)x]\,y_k^{(0)}y_l^{(0)}. \qquad (2.1.23)$$

Введем переменную $n = k-l$ и изменим порядок суммирования по k и n. Тогда

$$P_N(\nu) = \frac{1}{N^2}\left\{\sum_{k=0}^{N-1}\left[y_k^{(0)}\right]^2 + 2\sum_{n=1}^{N-1}\cos(2\pi nx)\sum_{k=0}^{N-1-n} y_k^{(0)}y_{k+n}^{(0)}\right\}. \qquad (2.1.24)$$

Полагая здесь

$$c_n = \frac{1}{N}\sum_{k=0}^{N-1-n} y_k^{(0)}y_{k+n}^{(0)}, \quad n = 0,\ 1,\ \ldots,\ N-1, \qquad (2.1.25)$$

получаем

$$P_N(\nu) = \frac{1}{N}\left[c_0 + 2\sum_{n=1}^{N-1}\cos(2\pi xn)\,c_n\right]. \qquad (2.1.26)$$

Величина c_0 представляет собой *выборочную дисперсию временного ряда*:

$$c_0 \equiv s^2 = \frac{1}{N}\sum_{k=0}^{N-1}\left[y_k^{(0)}\right]^2, \qquad (2.1.27)$$

а вся совокупность $\{c_n\}$, $n = 0,\ 1,\ 2,\ \ldots,\ N-1$ - *выборочную автоковариационную функцию.*

Выражению (26) можно придать более симметричный вид, если положить $c_{-n} = c_n$. Тогда находим окончательно

$$P_N(\nu) = \frac{1}{N}\sum_{n=-(N-1)}^{N-1}\cos(2\pi nx)\,c_n, \qquad (2.1.28)$$

$$|\nu| \leqslant \nu_c,\quad x = \nu\,\Delta t.$$

С помощью (П.2.4) можно показать, что (28) допускает обращение:

$$c_n = N\Delta t\int_{-\nu_c}^{\nu_c} e^{i\cdot 2\pi\nu\Delta t n}\,P_N(\nu)\,d\nu =$$

$$= 2N\Delta t\int_0^{\nu_c}\cos(2\pi\nu\Delta t n)\,P_N(\nu)\,d\nu. \qquad (2.1.29)$$

Таким образом, периодограмма (с точностью до постоянного множителя N^{-1}) и выборочная автоковариационная функция связаны парой преобразований Фурье.

§ 2.2. Теорема отсчетов

Высказанные в предыдущем параграфе соображения поясняют причину, по которой выполнение условия (2.1.21) позволяет однозначно разделить слагаемые в представлении $D_N(\nu)$ в виде суммы (2.1.20) и тем самым восстановить исходный временной процесс $f(t)$. Если шаг дискретизации Δt при наблюдении процесса выбран произвольно, без учета характера изменений $f(t)$, то может случиться так, что выполняется условие

$$F(\nu) = 0 \quad \text{при} \quad |\nu| > \nu_*, \tag{2.2.1}$$

но граничная частота ν_* превосходит частоту Найквиста $\nu_c = (2\Delta t)^{-1}$. Тогда, чтобы избежать неоднозначности при наложении отдельных спектров (см. рис. 2.1.2), следует уменьшить интервал Δt настолько, чтобы выполнялось неравенство $\nu_* \leqslant \nu_c$, т.е. шаг дискретизации Δt отвечал условию

$$\Delta t \leqslant \Delta t_* \equiv \frac{1}{2\nu_*} \quad , \tag{2.2.2}$$

гарантирующему вместе с (1) восстановление $f(t)$.

Функции $f(t)$, для которых выполняется условие (1), называются *функциями с ограниченным спектром*. Мы пришли к заключению, что для функций с ограниченным спектром существует шаг дискретизации $\Delta t_* = (2\nu_*)^{-1}$ такой, что совокупность дискретных значений $f(k\Delta t_*)$ содержит полную информацию о непрерывном процессе $f(t)$. Это утверждение составляет суть *теоремы отсчетов* (sampling theorem), важное значение которой для теории передачи сигналов было установлено Котельниковым (1933) и Шенноном (1948, 1949). Конкретное представление $f(t)$ через совокупность дискретных отсчетов

$$f(t) = \sum_{k=-\infty}^{\infty} \frac{\sin[2\pi\nu_*(t-k\Delta t_*)]}{\nu_*(t-k\Delta t_*)} f(k\Delta t_*) \tag{2.2.3}$$

было известно в математике задолго до появления первых работ по теории информации (см. Уиттекер, 1935). Приведем доказательство формулы (3).

Предположим, что выполнено условие (1). Тогда общее представление интеграла Фурье (1.3.2) принимает вид

$$f(t) = \int_{-\nu_*}^{\nu_*} e^{i\cdot 2\pi\nu t} F(\nu)\,d\nu \,. \qquad (2.2.4)$$

Функцию $F(\nu)$ на конечном промежутке $[-\nu_*,\ \nu_*]$ можно разложить в *ряд Фурье*:

$$F(\nu) = \sum_{k=-\infty}^{\infty} e^{i\cdot\pi k\nu/\nu_*} B_k, \qquad (2.2.5)$$

где в соответствии с (1.4.11) коэффициенты

$$B_k = \frac{1}{2\nu_*} \int_{-\nu_*}^{\nu_*} e^{-i\cdot\pi k\nu/\nu_*} F(\nu)\,d\nu \,. \qquad (2.2.6)$$

Сравнение (4) и (6) показывает, что

$$B_k = \frac{1}{2\nu_*} f\left(-k\,\frac{1}{2\nu_*}\right) = \Delta t_* f(-k\,\Delta t_*). \qquad (2.2.7)$$

Подставляя (5) в (4) и учитывая соотношение (7), находим окончательно:

$$f(t) = \int_{-\nu_*}^{\nu_*} e^{i\cdot 2\pi\nu t}\,d\nu \sum_{-\infty}^{\infty} B_k e^{i\pi k(\nu/\nu_*)} = \qquad (2.2.8)$$

$$= \sum_{-\infty}^{\infty} \Delta t_* f(-k\Delta t_*)\ \frac{2\sin[2\pi\nu_*(t+k\Delta t_*)]}{t+k\Delta t_*}\,,$$

что совпадает с (3).

Обратим внимание на то обстоятельство, что для восстановления $f(t)$ требуется бесконечное число отсчетов. Однако если функция $f(t)$ имеет не только ограниченный спектр, но и ограниченную интервалом T протяженность во времени (точнее, пренебрежимо мала вне этого интервала), то она полностью задана $2T\nu_*$ отсчетами. В качестве таковых могут быть взяты не только значения $f(t)$ через равноотстоящие интервалы времени, но и, например, значения ее первых производных (Шеннон, 1949).

Весьма полезно обсудить несколько различных интерпретаций простой формулы (2), связывающей максимальную частоту в спектре процесса ν_* и требуемый для его восстановления шаг дискретизации Δt.

Заметим прежде всего, что спектр $f(t)$ не обязательно должен быть отличен от нуля на малых частотах; важно лишь, чтобы он был ограничен в области высоких частот.

Далее, из (2) следует, что критическим условием восстановления является равенство

$$\nu_s = 2\nu_* , \qquad (2.2.9)$$

где

$$\nu_s = \frac{1}{\Delta t} = 2\nu_c \qquad (2.2.10)$$

- *частота отсчетов* (sampling frequency). В форме (9) обсуждаемое условие применяется не менее часто, чем в форме (2): для однозначного восстановления непрерывного процесса с максимальной частотой спектра ν_* необходима частота отсчетов ν_s, по крайней мере вдвое превосходящая ν_*. Заметим, что при $\nu_s > 2\nu_*$ процесс с ограниченным спектром можно представить только через предшествующие отсчеты (Вайнштейн и Зубаков, 1960; Папулис, 1985).

Максимальной частоте спектра ν_* отвечает минимальное характерное время изменений процесса $\tau_* = (\nu_*)^{-1}$. Поэтому (2) можно записать в виде

$$\Delta t \leqslant \frac{\tau_*}{2} , \qquad (2.2.11)$$

который показывает, что для восстановления процесса необходимо на минимальном характерном времени изменений процесса произвести не менее двух отсчетов.

Наконец, следует иметь в виду, что наблюдения, выполненные с временным шагом Δt, содержат информацию только о частотах спектра, не превосходящих частоту Найквиста. Допускаемое иногда пренебрежение этим обстоятельством потребовало специального напоминания (Курц, 1983).

Трудно ожидать, что теоретическое условие (1) будет строго выполняться на практике. Если говорить о полезной информации - сигнале, - то соответствующий спектр всегда

ограничен ввиду отличия от нуля постоянной времени регистрирующего прибора (при наблюдении собственно временного процесса) или конечности разрешающей способности оптики (при восстановлении изображения). Однако спектральная плотность шума распределена обычно в значительно более широкой области частот, чем та область, которая несет полезную информацию. Поэтому прежде чем производить дискретизацию процесса, следует пропустить данные через фильтр низких частот, отсекающий высокочастотный «хвост» шума (см. Бролт и Уайт, 1971). В простейшем случае такой фильтр имеет прямоугольную форму, причем верхняя граница полосы частот ν_* подбирается так, что вблизи нее отношение сигнала к шуму приблизительно равно 1. Более строгая постановка задачи об оптимальной фильтрации обсуждается в гл. 3.

Пример 2.2.1. В связи с разработкой цифрового способа записи и воспроизведения музыки возникла необходимость оценить объем требующейся для этого памяти компьютера. Найдем приблизительно этот объем при записи симфонического оркестра в течение 1 часа.

Как известно, ухо человека воспринимает звуковые колебания в диапазоне частот примерно от 20 Гц до 20 кГц. Полагая $\nu_* = 2 \cdot 10^4$ Гц находим из (9) частоту отсчетов, гарантирующую полное сохранение качества звучания: $\nu_S = 4 \cdot 10^4$ Гц. Это означает, что колебания амплитуды звуковой волны со временем должны регистрироваться с шагом $\Delta t \leq 2{,}5 \cdot 10^{-5}$с = 25 мкс. При указанном выше значении ν_S за 1 час будет зарегистрировано $1{,}44 \cdot 10^8$ отсчетов. Как свидетельствуют Мэттьюз и Пирс (1987), для качественного воспроизведения симфонической музыки необходима дискретизация амплитуды звука не менее чем на $2^{16} = 65\ 536$ уровней. Тогда часовая запись содержит около 10^{13} единиц информации (бит) или $\sim 10^{12}$ байт. Полученное значение иллюстрирует трудности, с которыми приходится сталкиваться в области цифровой музыки (астрономическая фотопластинка, полученная с широкоугольным телескопом, содержит $\sim 10^{10}$ байт информации).

§ 2.3. Статистические свойства периодограммы

Критерии значимости. Одну из классических задач мате-

матической статистики можно иллюстрировать следующим примером. Предположим, что произведено 600 бросаний кубической игральной кости, в результате чего зарегистрировано n_1 выпаданий 1, n_2 выпаданий 2 ,..., n_6 выпаданий 6 (очевидно, $n_1 + n_2 + \ldots + n_6 = 600$). Интуитивно ясно, что даже для идеальной кости реализация строго равномерного набора $n_1 = n_2 = \ldots = n_6 = 100$ маловероятна; почти всегда мы будем иметь дело с одной из возможных неравномерных совокупностей $\{n_k\}$, в которой существует по крайней мере одно максимальное значение. Пусть, например, значение 3 выпадало чаще других. Возникает вопрос: достаточно ли значимо отклонение n_3 от среднего для идеальной кости значения 100, чтобы можно было считать ее несимметричной?

Если бы в нашем распоряжении имелась другая, строго симметричная кость, на поставленный вопрос можно было бы ответить экспериментальным путем. Для этого нужно произвести много серий из 600 бросаний идеальной кости и оценить, сколь часто встречаются случаи, когда хотя бы одно из значений $n_1,\ldots, n_6$ по крайней мере столь же велико, как и наблюдавшееся для реальной кости значение n_3. Если частота таких случаев невелика, скажем, меньше $q = 0,01$, то наблюдавшееся отклонение n_3 следует признать значимым, а реальную кость заподозрить в несимметричности. В противном случае полагают, что наблюдаемое значение n_3 вполне совместимо с гипотезой о симметричности реальной кости, ибо аналогичные отклонения характерны и для идеальной кости.

Указанный эксперимент не обязательно производить в действительности - мы можем точно решить задачу о бросаниях идеальной кости аналитически и найти то критическое значение n_q, которое соответствует принятой для редких событий вероятности q, называемой *уровнем значимости*. Одно из двух указанных выше решений выбирается в зависимости от того, больше n_3 критического значения n_q или меньше его.

Подобно тому как наличие максимального из чисел $n_1,\ldots, n_6$ не означает еще, что кость соответственно несимметрична, выделение в заданном интервале частот наибольшего максимума периодограммы не означает, что обнаружена истинная (детерминированная) периодичность: в случайной реализации

ряда конечной длины всегда встретится последовательность, близкая к периодической (см. рис. 1.1.9). Поэтому для оценки значимости наблюдаемого на некоторой частоте ν' максимума периодограммы $P_N(\nu')$ следует рассмотреть случайный временной ряд, заведомо не содержащий никаких детерминированных периодических зависимостей, и найти вероятность q' того, что периодограмма случайного ряда на какой-либо частоте превзойдет $P_N(\nu')$. Если $q' < q$ - некоторого принятого уровня значимости, - то максимум считается значимым, в противном случае делается вывод об отсутствии указаний на реальную периодичность.

Нужно сказать, что иногда оценке значимости максимума периодограммы придают неоправданно большой вес, забывая о предварительной «чистке» периодограммы от ложных зависимостей, внесенных дискретностью наблюдений. Именно последняя задача ставит основные проблемы при анализе временных рядов в астрономии; после ее решения нахождение необходимых статистических оценок, как мы увидим ниже, производится достаточно просто.

Периодограмма случайного ряда. В соответствии со сказанным рассмотрим стационарный случайный временной ряд $\{\xi_k\}$, $k = 0, \pm1, \ldots$, отвечающий равномерно распределенным во времени моментам наблюдения $\{t_k\}$ с шагом Δt. Предположим, что среднее значение $\langle\xi_k\rangle = 0$, дисперсия $\langle\xi_k^2\rangle = \sigma^2$, и введем автоковариационную функцию

$$\psi_n = \langle\xi_n\xi_{k+n}\rangle, \quad n = 0, \pm1, \pm2, \ldots \qquad (2.3.1)$$

аналогично определению $\psi(\tau)$ для непрерывных процессов (1.5.4). Дискретной последовательности $\{\psi_n\}$ соответствует спектральная плотность (ср. с (1.5.13))

$$g(\nu) = \Delta t \sum_{n=-\infty}^{\infty} e^{-i\cdot 2\pi\nu\Delta t n}\,\psi_n = \Delta t \sum_{n=-\infty}^{\infty} \cos(2\pi\nu\Delta t n)\psi_n =$$

$$= \sigma^2\Delta t \left[1 + 2\sum_{n=1}^{\infty} \cos(2\pi\nu\Delta t n)\rho_n\right], \qquad (2.3.2)$$

где

$$\rho_n = \frac{\psi_n}{\sigma^2}, \quad \rho_0 = 1, \quad \rho_{-n} = \rho_n, \quad n = 0, 1, 2\ldots \qquad (2.3.3)$$

- автокорреляционная функция последовательности $\{\xi_k\}$, а частота заключена в промежутке $[-\nu_c, +\nu_c]$. Обратим внимание, что мы ограничились в (2) так называемым *главным частотным интервалом*, поскольку для равномерного ряда он содержит полную спектральную информацию. Как и (1.5.13), преобразование (2) допускает обращение

$$\psi_n = \int_{-\nu_c}^{\nu_c} e^{i\cdot 2\pi\nu\Delta t n}\, g(\nu)d\nu, \quad n = 0, \pm 1, \ldots \qquad (2.3.4)$$

В частности, при $n = 0$ отсюда следует аналог теоремы Рэлея:

$$\sigma^2 = \int_{-\nu_c}^{\nu_c} g(\nu)d\nu. \qquad (2.3.5)$$

Допустим теперь (гипотеза H_0), что изучаемый временной ряд $\{y_k^{(0)}\}$ представляет собой совокупность N последовательных значений, выбранных из чисто случайного ряда бесконечной длины $\{\xi_k\}$. Общее определение периодограммы (2.1.22) сводится в данном случае к следующему:

$$P_N(\nu) = U_N^2(\nu) + V_N^2(\nu), \quad -\nu_c \le \nu \le \nu_c,$$

$$U_N(\nu) = \frac{1}{N}\sum_{k=0}^{N-1} \cos(2\pi\nu\Delta t k)\,\xi_k, \qquad (2.3.6)$$

$$V_N(\nu) = \frac{1}{N}\sum_{k=0}^{N-1} \sin(2\pi\nu\Delta t k)\,\xi_k.$$

Для случайных временных рядов помимо периодограммы $P_N(\nu)$ целесообразно рассмотреть функцию

$$g_N(\nu) = N\Delta t P_N(\nu), \qquad (2.3.7)$$

которую именуют выборочной спектральной плотностью. Эта функция связана с выборочной автоковариационной функцией c_n формулами (2.1.25), (2.1.28) и (2.1.29), которые теперь принимают вид

$$c_n = \frac{1}{N}\sum_{k=0}^{N-1-n} \xi_k\xi_{k+n}, \quad c_{-n} = c_n, \quad n = 0, 1, \ldots, N-1, \qquad (2.3.8)$$

$$g_N(\nu) = \Delta t \sum_{n=-N+1}^{N-1} \cos(2\pi\nu\Delta t n)\,c_n, \qquad (2.3.9)$$

$$c_n = \int_{-\nu_c}^{\nu_c} e^{i\cdot 2\pi\nu\Delta t n}\, g_N(\nu)\, d\nu. \qquad (2.3.10)$$

Усредним соотношения (8) и (9) по вероятностному ансамблю. Принимая во внимание (1), находим

$$\langle c_n \rangle = \left(1 - \frac{|n|}{N}\right)\psi_n, \qquad (2.3.11)$$

так что

$$\langle g_N(\nu)\rangle = \Delta t \sum_{n=-N+1}^{N-1} \cos(2\pi\nu\Delta t n)\left(1-\frac{|n|}{N}\right)\psi_n. \qquad (2.3.12)$$

Как показывает сравнение формул (2) и (12),

$$\lim_{N\to\infty} \langle g_N(\nu)\rangle = g(\nu), \qquad (2.3.13)$$

т.е. средняя выборочная спектральная плотность случайного временного ряда стремится при увеличении его длины к истинной спектральной плотности этого ряда. Менее строгое обоснование этого важного результата уже было дано при обсуждении (1.6.4).

Нахождение среднего значения случайной функции $g_N(\nu)$ составляет лишь наиболее простую часть задачи; нам необходимо знать вероятностное распределение значений $g_N(\nu)$ или, что эквивалентно, периодограммы $P_N(\nu)$. Рассмотрим это распределение для того случая, когда ряд $\{\xi_k\}$ является белым шумом (см. пример 1.5.1). Тогда $\psi_n = 0$ при $n = 1, 2, \ldots$, спектральная плотность (2) ряда постоянна:

$$g(\nu) = \sigma^2\Delta t, \quad -\nu_c \leqslant \nu \leqslant \nu_c, \qquad (2.3.14)$$

а общее соотношение (13) принимает вид

$$\lim_{N\to\infty} \langle g_N(\nu)\rangle = \sigma^2\Delta t\ . \qquad (2.3.15)$$

Найдем сначала функции распределения случайных величин $U_N(\nu)$ и $V_N(\nu)$, входящих в (6). Их средние значения ввиду $\langle\xi_k\rangle = 0$ также равны нулю. Дисперсия $U_N(\nu)$ равна

$$\langle U_N^2(\nu)\rangle = \frac{\sigma^2}{N^2}\sum_{k=0}^{N-1}\cos^2(2\pi\nu\Delta t k), \qquad (2.3.16)$$

поскольку ξ_k и ξ_l некоррелированы при $k \neq l$. Предположим для определенности, что $N = 2J$ - четное число, и будем рассматривать все оценки для дискретной сетки частот

$$\nu_j = \frac{j}{N\Delta t} = \frac{j}{J}\nu_c, \quad j = -(J-1), \ldots, -1, 0, 1, \ldots, J, \qquad (2.3.17)$$

или даже , вследствие симметрии периодограммы, лишь при $0 \leqslant j \leqslant J$. Полагая $\nu = \nu_j$ в (16) и учитывая формулу (П.1.5), находим:

$$\langle U_N^2(\nu_j)\rangle = \begin{cases} \dfrac{\sigma^2}{2N}, & j = 1, 2, \ldots, J-1; \\ \dfrac{\sigma^2}{N}, & j = 0, J. \end{cases} \qquad (2.3.18)$$

Аналогично

$$\langle V_N^2(\nu_j)\rangle = \begin{cases} \dfrac{\sigma^2}{2N}, & j = 1, 2, \ldots, J-1; \\ 0, & j = 0, J, \end{cases} \qquad (2.3.19)$$

так что из (6), (18) и (19) следует, что

$$\langle P_N(\nu_j)\rangle = \frac{\sigma^2}{N}, \quad \langle g_N(\nu_j)\rangle = \sigma^2\Delta t, \qquad (2.3.20)$$
$$j = 0, 1, \ldots, J,$$

- для белого шума (15) выполняется не только асимптотически, но и при конечных значениях N.

Далее с помощью формул (П.1.1) и (П.1.2) нетрудно показать, что

$$\langle U_N(\nu_j)U_N(\nu_l)\rangle = \langle V_N(\nu_j)V_N(\nu_l)\rangle = 0, \quad j \neq l; \qquad (2.3.21)$$
$$\langle U_N(\nu_j)V_N(\nu_l)\rangle = 0, \quad j, l = 0, 1, \ldots, J.$$

Таким образом, величины $U_N(\nu_j)$, относящиеся к различным частотам, не коррелированы между собой; то же справедливо и для $V_N(\nu_j)$. Вместе с тем $U_N(\nu_j)$ и $V_N(\nu_j)$ не коррелированы даже в том случае, если они относятся к одной и той же частоте.

Для нахождения функции распределения U_N и V_N нам придется еще сузить гипотезу H_0 и предположить, что $\{\xi_k\}$ имеют

нормальное распределение, т.е. рассматриваемый временной ряд представляет собой гауссовский белый шум. Как известно (см. Крамер, 1946, п.17.3), сумма любого числа независимых нормально распределенных случайных величин сама распределена нормально. Поэтому при $j \neq 0, J$ величины $U_N(\nu_j)$ и $V_N(\nu_j)$ распределены нормально со средним 0 и дисперсией $\sigma^2/(2N)$. Также известно, что некоррелированные нормальные величины являются и независимыми. Принимая во внимание равенства (6) и (20), мы приходим к выводу, что случайная величина

$$\eta_j \equiv \frac{P_N(\nu_j)}{\langle P_N(\nu)\rangle} = \frac{g_N(\nu_j)}{\sigma^2}, \quad j = 0, 1, \ldots, J, \tag{2.3.22}$$

допускает при $j \neq 0, J$ представление

$$\eta_j = \frac{1}{2}(n_1^2 + n_2^2), \quad j = 1, 2, \ldots, J-1, \tag{2.3.23}$$

где n_1 и n_2 - независимые нормальные случайные величины со средним 0 и дисперсией 1. Нетрудно показать (Крамер, 1946, п.18.1), что плотность распределения величины (23) есть экспоненциальная функция e^{-x}, $x \geqslant 0$. При $j = 0, J$ величина $U_N(\nu_j)$ нормально распределена со средним 0 и дисперсией σ^2/N, а $V_N(\nu_j) \equiv 0$, так что η_j допускает представление

$$\eta_j = n_3^2, \quad j = 0, J, \tag{2.3.24}$$

где n_3 - нормальна с параметрами (0,1). Плотность распределения величины (24) равна $(2\pi x)^{-1/2} e^{-x/2}$, $x \geqslant 0$.

Таким образом, значения периодограммы $P_N(\nu_j)$ и выборочной спектральной плотности $g_N(\nu_j)$, нормированные к соответствующим средним согласно (22), имеют плотность распределения

$$s_j(x) = \begin{cases} e^{-x}, & j = 1, 2, \ldots, J-1; \\ (2\pi x)^{-1/2} e^{-x/2}, & j = 0, J; \end{cases} \tag{2.3.25}$$

где $J = N/2$ и $0 \leqslant x < \infty$. Этот фундаментальный результат был получен Шустером (1898). В математической литературе его чаще формулируют следующим образом: случайная

величина η_j имеет при $j = 0, J$ распределение χ^2 с одной степенью свободы, а при $j \neq 0, J$ величина $2\eta_j$ подчиняется распределению χ^2 с двумя степенями свободы.

Иногда вместо η_j в качестве статистики употребляется величина $\sqrt{\eta_j}$. Плотность распределения последней величины внутри главного частотного интервала есть $2xe^{-x^2}$, $x \geqslant 0$; эта функция называется *распределением Рэлея*.

Можно показать, что в случае $N \gg 1$, как это чаще всего и бывает на практике, распределение Шустера (25) справедливо не только для гауссовского, но и для почти произвольного белого шума. При этом нормальность U_N и V_N обеспечивается центральной предельной теоремой, а остальные рассуждения сохраняют силу.

Важность распределения Шустера для понимания эффективности спектрального анализа временных рядов определяется тем обстоятельством, что значения η_j одинаково распределены независимо от длины ряда N. Отсюда, в частности, следует, что дисперсия η_j не уменьшается с ростом N:

$$D(\eta_j) = \langle\eta_j^2\rangle - \langle\eta_j\rangle^2 = \begin{cases} 1, & j = 1, 2, \ldots, J-1; \\ 2, & j = 0, J. \end{cases} \qquad (2.3.26)$$

Если учесть, что значения η_j, разделенные интервалом частот порядка спектрального разрешения $\delta\nu = (N\Delta t)^{-1}$ независимы, становится понятной причина, по которой периодограммы (выборочные спектральные плотности) экспериментальных данных имеют столь «изрезанный» вид, не становящийся более гладким при сколь угодно значительном увеличении длины ряда. Другими словами, выборочная спектральная плотность не обладает свойством эргодичности, для нее среднее по времени не совпадает со средним по вероятностному ансамблю:

$$\lim_{N\to\infty} g_N(\nu) \neq \langle g_N(\nu)\rangle. \qquad (2.3.27)$$

Согласно (26), стандартные отклонения величин $P_N(\nu)$ и $g_N(\nu)$ внутри главного частотного интервала равны соответствующим средним значениям:

$$D^{1/2}[P_N(\nu)] = \langle P_N(\nu)\rangle, \quad D^{1/2}[g_N(\nu)] = \langle g_N(\nu)\rangle. \qquad (2.3.28)$$

Эти соотношения справедливы для случайных рядов довольно

общего вида (Бартлетт, 1955; Кендалл и Стюарт, 1976). Следует обратить внимание на тот факт, что распределение Шустера убывает при $x \to \infty$ существенно медленнее, чем привычное для экспериментаторов нормальное распределение. Это означает, что при том же уровне значимости допустимы случайные флуктуации периодограммы белого шума, значительно превосходящие среднее значение (20). Ограничимся внутренними точками главного частотного интервала и рассмотрим простой пример, иллюстрирующий различное поведение статистик для нормального и экспоненциального распределений. Для первого из них вероятность односторонней случайной флуктуации, превосходящей три стандартных отклонения, равна 0,00135. Малость этой вероятности и обусловливает широкое распространение «правила 3σ» при оценке значимости одного измерения нормальной величины. Трем стандартным отклонениям экспоненциальной величины отвечает уровень значимости $e^{-3} \approx 0,050$, а для достижения того же уровня $q = 0,00135$ нужно применить «правило 6,6σ».

Значимость максимумов периодограммы. Распределение (25) определяет поведение периодограммы на какой-либо фиксированной частоте. Однако нередко значение возможного периода изучаемого процесса неизвестно заранее. В этом случае необходимо оценить значимость наибольшего из отсчетов периодограммы в главном частотном интервале. Полагая $N > 2$ и ограничиваясь внутренними точками этого интервала, мы можем считать, что в нем имеется $J-1 = N/2-1$ независимых значений нормированной периодограммы (спектральной плотности) $\eta_1, \eta_2, \ldots, \eta_{J-1}$. Нужно найти вероятность того, что наибольшая из этих случайных величин ζ превосходит некоторое x, т. е.

$$\Pr\{\zeta > x\} \equiv Q_N^{(W)}(x), \qquad \zeta = \max(\eta_1, \ldots, \eta_{J-1}). \tag{2.3.29}$$

Вероятность того, что любая из η_j не превосходит x, равна

$$S(x) = \int_0^x s_j(x')dx' = 1 - e^{-x}, \quad x \geqslant 0. \tag{2.3.30}$$

Далее, вероятность того, что все случайные величины $\eta_1, \ldots, \zeta_{J-1}$ не превосходят x, равна $(1-e^{x})^{J-1}$, а вероятность того, что хотя бы одна из η_j больше x, есть

$$Q_N^{(W)}(x) = 1 - (1-e^{-x})^{N/2-1}, \quad 0 \leqslant x < \infty. \tag{2.3.31}$$

Это выражение, полученное Уокером (1914), и описывает распределение наибольшего из значений периодограммы или выборочной спектральной плотности в главном частотном интервале. Мы будем называть его *распределением Уокера.*

Уровню значимости q соответствует отклонение

$$X_{N,q}^{(W)} \equiv \frac{P_N^{max}(\nu)}{\langle P_N(\nu)\rangle} = \frac{g_N^{max}(\nu)}{\sigma^2}, \tag{2.3.32}$$

являющееся решением уравнения

$$Q_N^{(W)}(x) = q \tag{2.3.33}$$

относительно x, т.е.

$$X_{N,q}^{(W)} = -\ln\left[1 - (1-q)^{\frac{2}{N-2}}\right]. \tag{2.3.34}$$

Итак, при заданных длине ряда N и уровне значимости q лишь отклонения нормированной периодограммы (или спектральной плотности), превосходящие $X_{N,q}$, могут считаться значимыми.

До сих пор мы не учитывали, что среднее значение периодограммы $\langle P_N(\nu)\rangle$, которое необходимо для нахождения нормированного отклонения (22) и, следовательно, применения распределения Уокера (31), строго говоря, остается неизвестным. На опыте мы можем оценить лишь выборочное среднее значение

$$\overline{P_N}(\nu) = \frac{1}{J-1}\sum_{j=1}^{J-1} P_N(\nu_j), \tag{2.3.35}$$

которое представляет собой случайную величину, распределенную вблизи среднего по ансамблю $\langle P_N(\nu)\rangle = \sigma^2/N$. При достаточно большой длине ряда N следует ожидать, что отличие истинного и выборочного средних невелико, однако при неболь-

щих значениях N вместо (22) необходимо рассматривать случайные величины

$$\eta'_j = \frac{P_N(\nu_j)}{\bar{P}_N(\nu)} = \frac{g_N(\nu_j)}{\bar{g}_N(\nu)}, \quad j = 1, 2, \ldots, J-1, \qquad (2.3.36)$$

а вместо (29) - случайную величину

$$\zeta' = \max(\eta'_1, \eta'_2, \ldots, \eta'_{J-1}). \qquad (2.3.37)$$

Соответствующее вероятностное распределение, заменяющее распределение Уокера (31), было найдено Фишером (1929). *Распределение Фишера* имеет вид

$$\Pr\{\zeta' > (J-1)t\} \equiv Q_N^{(F)}[(J-1)t] =$$

$$= \sum_{k=1}^{r} (-1)^{k-1} C_{J-1}^{k} (1-kt)^{J-2}, \qquad (2.3.38)$$

где r - наибольшее целое, не превосходящее $J-2$, для которого $1-rt \geqslant 0$, а C_n^m - число сочетаний из n элементов по m.

Для нас основной интерес представляет оценка величины $X_{N,q}^{(F)}$, являющейся решением уравнения

$$Q_N^{(F)}(x) = q \qquad (2.3.39)$$

относительно x, т.е. более точного критического отклонения, заменяющего (34). Соответствующие численные расчеты были выполнены Дэвисом (1941). Кирк и др. (1979) дают таблицу значений $X_{N,q}^{(F)}$ при $q = 0,05$ и $q = 0,01$ согласно Фуллеру (1976). Эта таблица приводится в Приложении 5, а также иллюстрируется рис. 2.3.1. Заметим, что прямолинейные участки зависимости $X_{N,q}^{(F)}$ от N при $N \gg 1$ соответствуют (34), как этого и следовало ожидать из соображений, высказанных в начале данного пункта. При $q \ll 1$, $N \gg 1$ можно пользоваться простой приближенной формулой

$$X_{N,q} \simeq \ln \frac{N-2}{2q}, \qquad (2.3.40)$$

следующей из (34). Возрастание $X_{N,q}$ с ростом N обусловлено очевидной причиной: чем больше совокупность независимых реализаций, тем большие случайные отклонения являются вполне допустимыми при данном уровне значимости q.

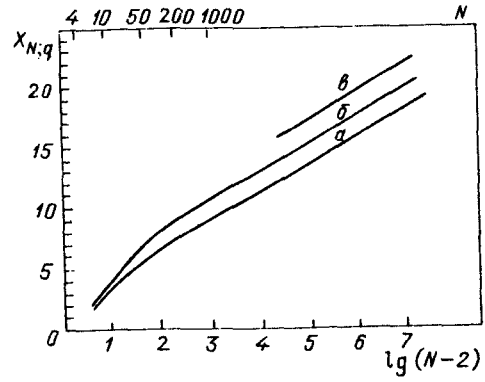


Рис. 2.3.1. Максимальное отношение $P_N(\nu)/\overline{P_N(\nu)}$ в зависимости от количества отсчетов N и уровня значимости q = 0,05 (а), 0,01 (б) и 0,00135 (в)

Пользуясь формулой (40), можно дать приближенную оценку амплитуды A гармонической составляющей $\tilde{f}(t) = A\cos(2\pi\nu_0 t-\varphi)$, которая еще обнаружима в отрезке случайного временного ряда с дисперсией σ^2. Как показывает (20), среднее значение периодограммы ряда равно $\langle P_N\rangle = \sigma^2/N$. С другой стороны, согласно (1.6.8) вклад гармонического колебания $\tilde{f}(t)$ в значение периодограммы на частоте ν_0 равен приблизительно $A^2/4$, так что отношение (22) есть $\eta \simeq A^2N/(4\sigma^2)$. При этом уровню значимости q отвечает неравенство $\eta \geqslant X_{N;q}$ или

$$\frac{A}{\sigma} \geqslant \left[\frac{4}{N}\ln\frac{N}{2q}\right]^{1/2}, \qquad (2.3.41)$$

где мы полагаем $q \ll 1$ и $N \geqslant 30$. Так, при $N = 50$ уровню $q = 0,05$ соответствует минимальная амплитуда обнаружимой гармоники $A_{min} \simeq 0,71\sigma$, а при $N = 200$ этому же значению уровня значимости отвечает $A_{min} \simeq 0,39\sigma$. Представляет интерес и обращение формулы (41), дающее уровень значимости обнаружения гармонической составляющей амплитуды A в случайном шуме с дисперсией σ^2:

$$q \simeq \frac{N}{2}\, e^{-N\left(\frac{A}{2\sigma}\right)^2}, \qquad (2.3.42)$$

где N - число отсчетов временного ряда. Следует ожидать, что два последних соотношения приближенно справедливы и для

рядов, не слишком сильно отличающихся от равномерного ряда. Необходимые в этом случае уточнения можно внести путем численных расчетов методом Монте-Карло.

Пример 2.3.1. Рассматривавшийся в § 1.1 ряд А включал в себя случайный компонент $\xi(t)$, периодограмма которого представлена на рис. 1.1.9. Наибольший из максимумов $P_N(\nu)$ расположен на частоте $\nu_* \simeq 0{,}228$ и равен $P_N(\nu_*) \simeq 0{,}021$. Можно ли считать этот максимум значимым на уровне $q = 0{,}01$?

Среднее значение периодограммы в диапазоне частот $0 < \nu < 0{,}5$ равно $\overline{P}_N(\nu) \approx 0{,}0046$, так что мы имеем реализацию случайной величины ζ', равную согласно (36) и (37) $\zeta'_* = 0{,}021/0{,}0046 \simeq 4{,}57$. Случаю $N = 50$, $q = 0{,}01$ соответствует по табл. П.5.1 отклонение $X_{50;\,0{,}01} \simeq 6{,}89$. Поэтому указанный максимум нельзя считать значимым на 1 %-м уровне. Из той же таблицы П.5.1 видно, что он незначим и на 5 %-м уровне. Для приближенной оценки уровня значимости, отвечающего $\zeta'_* = 4{,}57$, воспользуемся формулой (31); из нее следует $Q_{50}^{(W)}(4{,}57) \approx 0{,}22$, т.е. в 22 % случаев заведомо случайный ряд длиной $N \sim 50$ обнаружит максимумы периодограммы по крайней мере такие же большие, как и наблюдавшийся на самом деле. Таким образом, рассматриваемый ряд не содержит значимых свидетельств периодичности. Но для максимума на частоте $\hat{\nu}_0 \approx 0{,}370$, отвечающего истинному периодическому компоненту (см. рис. 1.1.2), $\zeta'_0 \approx 13$, и мы не ошибемся, если сочтем его значимым: $Q_{50}^{(W)}(13) \approx 5 \cdot 10^{-5}$.

Иногда при наблюдениях приходится оценивать значимость проявления периодичности с известным заранее значением периода (например, отыскиваются оптические колебания блеска двойной звезды с известным по рентгеновским данным периодом). Тогда при уровне значимости q отношение η высоты максимума периодограммы $A^2/4$ к среднему ее значению σ^2/N должно превосходить некоторое X_q, равное, ввиду экспоненциального характера распределения η, просто $-\ln q$. Поэтому в рассматриваемом случае минимальная амплитуда

$$A_{\min} \simeq 2\sigma \sqrt{\frac{-\ln q}{N}} \,. \qquad (2.3.43)$$

В заключение параграфа отметим, что наличие шумовой со-

ставляющей ограничивает не только амплитуду обнаружимого периодического сигнала, но и точность определения его частоты. Как уже неоднократно указывалось, ширина максимума периодограммы временного ряда длительности T составляет величину $\delta\nu_0 \simeq T^{-1}$, так что представляет интерес выразить стандартное отклонение неточности определения частоты вследствие шума $\delta\nu_n$ в единицах $\delta\nu_0$. Воспользовавшись результатами Ковача (1981), нетрудно показать, что

$$\frac{\delta\nu_n}{\delta\nu_0} \simeq \frac{\sigma/A}{\sqrt{N}}, \qquad (2.3.44)$$

где σ – стандартное отклонение шума, A - амплитуда сигнала, N- число отсчетов временного ряда. Есть основания полагать, что приведенная формула справедлива и для неравномерных рядов, не содержащих значительных перерывов в наблюдениях.

§ 2.4. Сглаживание периодограммы

Анализ какого-либо процесса на основании результатов измерений обычно имеет своей целью получение оценок неизвестных параметров или функций; при этом исследователь стремится найти *несмещенные* и *состоятельные оценки*. Первый из этих терминов указывает, что среднее значение оценки должно быть равно истинному (теоретическому) значению соответствующего параметра или функции, второй - что дисперсия оценки должна стремиться к нулю по мере роста объема выборки N, т.е. точность оценки возрастать при $N \to \infty$.

Соотношение (2.3.13) показывает, что выборочная спектральная плотность $g_N(\nu) \equiv N\Delta t P_N(\nu)$, рассматриваемая в качестве оценки истинной спектральной плотности процесса $g(\nu)$, является, вообще говоря, асимптотически несмещенной. При конечных значениях N смещение равно нулю лишь для белого шума. Обычно устранение смещения оценки не составляет особых трудностей. Значительно более серьезная проблема связана с тем, что $g_N(\nu)$ является несостоятельной оценкой $g(\nu)$, поскольку ее дисперсия не зависит, согласно (2.3.26), от длины временного ряда. Это означает, что относительная величина флуктуаций $g_N(\nu)$ и периодограммы $P_N(\nu)$ не убывает с

ростом N, т.е. указанные функции остаются в равной степени «изрезанными» и для коротких, и для сколь угодно длинных временных рядов. Ситуация здесь аналогична изрезанности гистограммы при слишком малом интервале разбиения аргумента (Бокс и Дженкинс, 1970).

В этой связи естественно попытаться каким-либо образом сгладить значения $g_N(\nu)$ и получить в результате состоятельную оценку $g(\nu)$. Например, можно просто усреднить соседние значения $g_N(\nu)$ в движущемся вдоль спектра интервале частот фиксированной ширины (при этом выборочный спектр как бы просматривается через «спектральное окно» прямоугольной формы). Другие подходы связаны с введением некоторых дополнительных операций на более ранней стадии, предшествующей вычислению выборочной спектральной плотности. К ним относится, в частности, весовая обработка ряда или его автоковариационной функции. Сглаживание спектра действительно приводит к состоятельным оценкам, однако при этом приходится искать компромисс между устойчивостью оценки спектра, смещением и спектральным разрешением.

Сглаживание временного окна. Тот факт, что наблюдениям доступна лишь часть $(\xi_0, \xi_1, \ldots, \xi_{N-1})$ бесконечной случайной последовательности $\{\xi_k\}$, формально эквивалентен умножению $\{\xi_k\}$ на прямоугольную функцию типа $r_T(t)$, определенной равенством (1.3.19). В § 1.6 сходная процедура было описана с помощью функции временного окна $w_N(t)$ и соответствующей ей функции спектрального окна $W_N(\nu)$. Резкий обрыв временной последовательности $(\xi_0, \ldots, \xi_{N-1})$ обусловлен, разумеется, лишь возможностями наблюдения, и поэтому интересно выяснить, к каким последствиям приведет сглаживание «углов» функции временного окна.

Основная цель такой процедуры заключается в повышении спектрального разрешения. Действительно, из (1.3.25) видно, что применение прямоугольного временного окна эквивалентно введению в спектральной области отклика $T\,\mathrm{sinc}(\nu T)$, обладающего довольно интенсивными боковыми максимумами (рис. 1.3.1). Сглаживанием прямоугольного окна можно достичь перераспределения энергии в основном в боковых максимумах, так что центральный максимум становится несколько

шире, но боковые лепестки сильно подавляются. Аналогичная процедура, называемая в оптике *аподизацией*, иногда применяется при астрономических наблюдениях, например, поисках двойных звезд.

Введем с этой целью в определения (2.3.6), (2.3.7) периодограммы и выборочной спектральной плотности некоторые весовые множители $\{w_k\}$, плавно убывающие к границам изучаемого временного интервала:

$$g_N^{(w)}(\nu) \equiv N\Delta t P_N^{(w)}(\nu) = \frac{\Delta t}{N}\left|F_N^{(w)}(\nu)\right|^2,$$
$$F_N^{(w)}(\nu) = \sum_{k=0}^{N-1} e^{-i\cdot 2\pi\nu\Delta t k}\, w_k \xi_k . \qquad (2.4.1)$$

Как и в (1.6.18), это эквивалентно дискретизации исходного процесса:

$$\xi_N(t) = \xi(t) w_N(t), \qquad (2.4.2)$$

где временное окно

$$w_N(t) = \frac{1}{N}\sum_{k=0}^{N-1} w_k \delta(t-k\Delta t). \qquad (2.4.3)$$

Соответствующее спектральное окно имеет теперь вид

$$W_N(\nu) = \frac{1}{N}\sum_{k=0}^{N-1} w_k e^{-i\cdot 2\pi\nu\Delta t k}, \quad -\nu_c \leqslant \nu \leqslant \nu_c, \qquad (2.4.4)$$

а основное уравнение (1.6.31), связывающее истинную спектральную плотность $g(\nu)$ и среднюю периодограмму $\langle P_N(\nu)\rangle$, сохраняет вид при расширенном определении периодограммы:

$$\langle P_N^{(w)}(\nu)\rangle = \int_{-\nu_c}^{\nu_c} |W_N(\nu-\nu')|^2 g(\nu')\,d\nu' \qquad (2.4.5)$$

или

$$\langle g_N^{(w)}(\nu)\rangle = \int_{-\nu_c}^{\nu_c} L_N(\nu-\nu') g(\nu')\,d\nu', \qquad (2.4.6)$$

где

$$L_N(\nu) = N\Delta t |W_N(\nu)|^2. \qquad (2.4.7)$$

Если определить весовые множители $\{w_k\}$ так, чтобы

выполнялось условие

$$\frac{1}{N}\sum_{k=0}^{N-1} w_k^2 = 1, \qquad (2.4.8)$$

то ядро $L_N(\nu)$ станет нормированным к 1:

$$\int_{-\nu_c}^{\nu_c} L_N(\nu)\,d\nu = 1. \qquad (2.4.9)$$

Поскольку $L_N(\nu)$ - периодическая функция с периодом $\Delta t^{-1} = 2\nu_c$, выполнение (9) обеспечивает равенство интегралов от $\langle g_N(\nu)\rangle$ и $g(\nu)$ по главному частотному интервалу (сохранение энергии).

Соотношение (6) показывает, что в случае достаточно гладкой спектральной плотности $g(\nu)$ подходящим выбором весов можно добиться, чтобы смещение выборочной спектральной плотности было невелико. Очевидно, для этого нужно сделать ядро $L_N(\nu)$ как можно более «узким», и тогда, вынося $g(\nu)$ в (6) за знак интеграла, получаем ввиду (9) $\langle g_N^{(w)}(\nu)\rangle \simeq g(\nu)$. Однако сглаживание временного окна не изменяет статистических свойств периодограммы при $N \gg 1$ (Бриллинджер, 1975, теорема 5.2.7), так что (1) по-прежнему не может служить состоятельной оценкой истинной спектральной плотности. Качественно этот вывод ясен на основании упоминавшейся выше аналогии между временным сглаживанием и оптической аподизацией.

Сглаживание корреляционного окна. Эффективный способ получения состоятельных оценок спектральной плотности основан на предварительном сглаживании совокупности (2.3.8) выборочных коэффициентов ковариации $\{c_n\}$, $|n| \leqslant N-1$. Эти коэффициенты задают при помощи (2.3.9) выборочную спектральную плотность $g_N(\nu)$, и естественно думать, что именно резкий обрыв известной из наблюдений последовательности $\{c_n\}$ вблизи границ $n = \pm(N-1)$ обусловливает нерегулярное поведение $g_N(\nu)$.

Для того, чтобы компенсировать упомянутый обрыв последовательности $\{c_n\}$, можно учитывать далекие корреляции все с меньшим весом по мере роста $|n|$, так что сглаженная спект-

ральная плотность $g_N^{(\lambda)}(\nu) \equiv N\Delta t P_N^{(\lambda)}(\nu)$ по определению равна

$$g_N^{(\lambda)}(\nu) = \Delta t \sum_{k=-(N-1)}^{N-1} \cos(2\pi\nu\Delta t n)\lambda_n c_n, \qquad (2.4.10)$$

где $\{\lambda_n\}$ - выбранные подходящим образом весовые коэффициенты - *корреляционное окно* по терминологии, введенной Блэкмэном и Тьюки (1959). Эта совокупность произвольна за исключением условия $\lambda_0 = 1$, необходимого для сохранения энергии в главном частотном интервале. Действительно, согласно (2.3.5) интеграл от $g(\nu)$ равен истинной дисперсии процесса σ^2, а интеграл от (10) при $\lambda_0 = 1$ равен

$$\int_{-\nu_c}^{\nu_c} g_N^{(\lambda)}(\nu)\,d\nu = \int_{-\nu_c}^{\nu_c} g_N(\nu)\,d\nu = s^2, \qquad (2.4.11)$$

где

$$s^2 = c_0 = \frac{1}{N}\sum_{k=0}^{N-1}\xi_k^2 \qquad (2.4.12)$$

- (смещенная) выборочная оценка дисперсии процесса, сходящаяся при $N \to \infty$ к σ^2.

Явное выражение сглаженной спектральной плотности $g_N^{(\lambda)}(\nu)$ через выборочную плотность $g_N(\nu)$ получается при подстановке в (10) выражения (2.3.10) для c_n. Оно имеет вид

$$g_N^{(\lambda)}(\nu) = \int_{-\nu_c}^{\nu_c} \Lambda_N(\nu,\nu')g_N(\nu')\,d\nu', \qquad (2.4.13)$$

где ядро

$$\Lambda_N(\nu,\nu') = \Delta t \sum_{-(N-1)}^{N-1} \cos(2\pi x n)\cos(2\pi x' n)\lambda_n =$$

$$= \frac{\Delta t}{2}\sum_{-(N-1)}^{N-1}\{\cos[2\pi(x-x')n] + \cos[2\pi(x+x')n]\}\lambda_n, \qquad (2.4.14)$$

а безразмерные переменные суть $x = \nu\Delta t$ и $x' = \nu'\Delta t$. Ввиду $\lambda_0 = 1$ выполняется условие нормировки

$$\int_{-\nu_c}^{\nu_c} \Lambda_N(\nu,\nu')\,d\nu = 1. \qquad (2.4.15)$$

Таким образом, введение корреляционного окна $\{\lambda_n\}$ эквивалентно сглаживанию $g_N(\nu)$ при помощи плотности спектрального окна $\Lambda_N(\nu,\nu')$.

На практике вычисление всех коэффициентов ковариации c_n вплоть до $n = N-1$ слишком сложно при большой длине временного ряда. Обычно корреляционное окно выбирают значительно более узким, так что $\lambda_n \equiv 0$ при $|n| > M$, где $M \ll N$.

Поскольку цель сглаживания заключается в получении состоятельной оценки спектральной плотности, основной интерес представляет нахождение дисперсии величины

$$\eta_j^{(\lambda)} = \frac{g_N^{(\lambda)}(\nu_j)}{\langle g_N^{(\lambda)}(\nu)\rangle} = \frac{P_N^{(\lambda)}(\nu_j)}{\langle P_N^{(\lambda)}(\nu)\rangle}, \quad j = 1, 2, \ldots, J-1, \qquad (2.4.16)$$

аналогичной (2.3.22). Можно показать (см. например, Т. Андерсон, 1971, теорема 9.3.4), что при $N \to \infty$, $M \to \infty$, $M/N \to 0$ дисперсия нормированной спектральной плотности равна

$$D\left[\eta_j^{(\lambda)}\right] \simeq \frac{M}{N}\,\lambda_*, \qquad (2.4.17)$$

где параметр

$$\lambda_* = \frac{1}{M}\sum_{n=-M}^{M}\lambda_n^2. \qquad (2.4.18)$$

Мы видим, что дисперсия $\eta_j^{(\lambda)}$ стремится к 0 при $N \to \infty$, и сглаженная плотность $g_N^{(\lambda)}(\nu)$ является асимптотически состоятельной оценкой истинной спектральной плотности $g(\nu)$. Поскольку $g_N^{(\lambda)}(\nu)$ является и асимптотически несмещенной, можно утверждать (Бриллинджер, 1975), что оценка $g_N^{(\lambda)}(\nu)$ состоятельна в среднеквадратичном:

$$\lim_{T\to\infty} \langle |g_N^{(\lambda)}(\nu) - g(\nu)|^2 \rangle = 0. \qquad (2.4.19)$$

Что касается функции распределения значений сглаженной плотности, то при тех же условиях, которые приводят к (17), случайная величина

$$\zeta = \sqrt{\frac{M\lambda_*}{N}}\left[\eta_j^{(\lambda)} - 1\right] \qquad (2.4.20)$$

нормальна со средним 0 и дисперсией 1 (Андерсон, 1971). Это

позволяет использовать при оценке уровней значимости выбросов сглаженной периодограммы хорошо известные распределения порядковых статистик нормальной величины (Большев, Смирнов, 1965). При умеренных значениях N следует исходить из того, что $f\eta_j^{(\lambda)}$ подчиняется χ^2-распределению с числом степеней свободы $f \approx 2N/(M\lambda_*)$. Подробное обсуждение этого результата, полученного Бартлеттом (1950), содержится в книге Дженкинса и Ваттса (1969).

Все приведенные результаты, касающиеся процедуры сглаживания спектральной плотности, достаточно наглядны. Если во временной области исходный ряд длины $N\Delta t$ подразделяется на отдельные более короткие ряды длины $M\Delta t$, то в спектральной области разрешение падает от $\delta\nu \simeq (N\Delta t)^{-1}$ до $\Delta\nu \simeq (M\Delta t)^{-1}$. Это эквивалентно усреднению $g_N(\nu)$ при помощи спектрального окна шириной $\sim\Delta\nu$, причем окно охватывает приблизительно $\Delta\nu/\delta\nu \approx N/M$ независимых значений $g_N(\nu)$. Разумеется, дисперсия усредненной спектральной плотности $g_N^{(\lambda)}(\nu)$ уменьшается при этом в N/M раз, т.е. становится порядка M/N (см. (17) и (2.3.26)). Множитель λ_* порядка 1 учитывает точную форму спектрального окна. Как известно, суммирование случайных величин, подчиняющихся χ^2-распределению, вновь приводит к χ^2-величинам с возросшим соответственно числом степеней свободы. Отсюда ясно, что при небольшом числе слагаемых N/M сглаженная плотность имеет распределение χ^2 с $f \simeq 2(N/M)$ степенями свободы, а при $N/M \gg 1$ это распределение вследствие центральной предельной теоремы приближается к нормальному распределению.

Пример 2.4.1. Прямоугольное окно. Выберем весовые множители $\{\lambda_n\}$ следующим образом:

$$\lambda_n = \begin{cases} 1, & |n| \le M, \\ 0, & |n| > M. \end{cases} \qquad (2.4.21)$$

Сглаженная периодограмма вычисляется при этом по формуле

$$g_N^{(\lambda)}(\nu) = \Delta t \sum_{n=-M}^{M} \cos(2\pi\nu\Delta t n) c_n. \qquad (2.4.22)$$

Обозначая, как обычно, $x = \nu\Delta t$ и $x' = \nu'\Delta t$, находим из (14) и (П.1.7) плотность спектрального окна

$$\Lambda_N(\nu,\nu') = \Delta t \sum_{-M}^{M} \cos(2\pi x n)\cos(2\pi x' n) =$$

$$= \frac{2M+1}{2}\,\Delta t\left[\phi_{2M+1}(x-x') + \phi_{2M+1}(x+x')\right], \qquad (2.4.23)$$

где $\phi_N(x) = \sin(\pi N x)/N\sin(\pi x)$. Выполнение условия нормировки (15) обеспечивается здесь соотношением (П.3.12).

Свойства $\Lambda_N(\nu,\nu')$ легко усмотреть из обсуждения характеристик функции $\phi_N(x)$ в § 2.1. Обратим внимание на то, что выражение (23) может принимать и отрицательные значения. Параметр $\lambda_* \simeq 2$, так что дисперсия нормированной спектральной плотности равна, согласно (17),

$$D\left[\eta_j^{(\lambda)}\right] \simeq 2M/N. \qquad (2.4.24)$$

Прямоугольному окну отвечает довольно высокое разрешение, однако его использование неудовлетворительно как в связи с возможностью получения отрицательных значений оценки $g_N^{(\lambda)}(\nu)$, так и по причине значительных боковых лепестков (sidelobes) спектрального окна, приводящих к так называемой утечке (leakage) энергии из областей с высокой спектральной плотностью в соседние участки спектра.

Пример 2.4.2. Окно Бартлетта. Одна из наиболее часто применяющихся форм корреляционного окна введена Бартлеттом (1950, 1955):

$$\lambda_n = \begin{cases} 1 - \dfrac{|n|}{M}, & |n| \le M, \\ 0, & |n| > M. \end{cases} \qquad (2.4.25)$$

По вполне понятной причине это окно называют иногда треугольным.

Вычисление сглаженной спектральной плотности производится в данном случае по формуле

$$g_N^{(\lambda)}(\nu) = \Delta t \sum_{-M}^{M} \cos(2\pi\nu\Delta t n)\left(1-\frac{|n|}{M}\right)c_n, \qquad (2.4.26)$$

где c_n определяется выражением (2.3.8). Как легко видеть, плотность спектрального окна Бартлетта равна

$$\Lambda_N(\nu,\nu') = \frac{M\Delta t}{2}\left[\phi_M^2(x-x') + \phi_M^2(x+x')\right], \qquad (2.4.27)$$

т.е. суперпозиции ядер Фейера (см. нормировку (П.3.13)),

Окно Бартлетта допускает весьма простую интерпретацию (Дженкинс и Ваттс, 1969). Оказывается, его введение эквивалентно разбиению исходного ряда длины N на $r = N/M$ рядов длины M (r предполагается целым), вычислению спектральной плотности для каждого из рядов и усреднению этих частных плотностей:

$$g_N^{(\lambda)}(\nu) = \frac{1}{r}\sum_{j=1}^{r} g_{N/r}^{(j)}(\nu). \qquad (2.4.28)$$

Как и следовало ожидать, с ростом r гладкость оценки спектральной плотности возрастает, однако при этом растет и ее смещение. По этой причине выбор M должен производиться оптимальным образом.

Ввиду (II.1.8) параметр λ_* для окна Бартлетта при $M \gg 1$ приблизительно равен 2/3, так что

$$D\left[\eta_j^{(\lambda)}\right] \simeq \frac{2}{3r} = \frac{2}{3}\frac{N}{M}. \qquad (2.4.29)$$

Число степеней свободы $f = 3r$, величина $3r\eta_j^{(\lambda)}$ распределена как χ^2_{3r}, а при $r \gg 1$ величина $[\eta_j^{(\lambda)} - 1](1,5r)^{-1/2}$ нормальна со средним 0 и дисперсией 1.

Заключительные замечания. Мы уже отмечали, что выбор корреляционного окна $\{\lambda_n\}$ во многом произволен. Помимо описанных выше детально изучались окна Парзена, Блэкмэна и Тьюки, Даниэля, Хемминга и др. (см. Дженкинс и Ваттс, 1969; Андерсон, 1971). Проблема оптимального сглаживания детально рассматривается в монографии Журбенко (1987). Для конкретной задачи выбор окна определяется относительным значением, которое придается гладкости спектра, ее малому смещению, спектральному разрешению и утечке в боковые лепестки. С целью по возможности удовлетворить этим противоречивым требованиям применяются и способы комбинированного сглаживания с использованием временного и корреляционного взвешивания (Нэтолл и Картер, 1982).

Таким образом, применение спектральных окон обеспечивает достаточную гибкость процедуры оценивания спектральной плотности. Но с этим же обстоятельством связан и общий недостаток всего описанного подхода: он во многом носит рецептурный характер. По мнению Джейнса (1982), этот подход

полезен на предварительном, поисковом этапе анализа проблемы.

§ 2.5. Спектральное оценивание с помощью метода максимума энтропии

В настоящем параграфе мы познакомимся с теоретико-информационными методами восстановления изображения на примере конкретной задачи нахождения оценки спектральной плотности $g(\nu)$ случайного процесса по результатам его измерений $(y_0, y_1, \ldots, y_{N-1})$. Идеи, на которых основан описываемый подход, являются, по сути дела, общими как для анализа одномерных временных и пространственных рядов, так и для восстановления двумерных изображений при размытии заданной аппаратной функцией в присутствии шума.

Метод максимальной энтропии (Берг, 1967), явившийся первым значительным шагом на пути применения информационных критериев, выбирает из множества согласующихся с данными наблюдений оценок неизвестного оригинала ту оценку, которая обладает наибольшим статистическим весом (энтропией). Как нетрудно заметить, истоки указанного подхода связаны с вероятностной трактовкой понятий статистической физики Л. Больцманом (1872), дальнейшим развитием этих идей П. и Т. Эренфестами (1912) и работами К. Шеннона (1948) по теории передачи информации. Другие вероятностные критерии, например, методы *максимальной информации* и *максимального правдоподобия* исходят из требований, несколько отличных от указанного выше, однако все эти методы глубоко связаны между собой.

Характерная черта информационных методов анализа заключается в нелинейности восстанавливающего алгоритма, что обусловливает и необходимость выполнения значительного объема вычислений (обычно - некоторого итеративного процесса). Вместе с тем выяснилось, что нелинейность задачи является формальным отражением принципиальных преимуществ информационого подхода по сравнению с применявшимися ранее линейными методами (например, оптимальной фильтрацией Вине-

ра). Именно, информационный подход 1) позволяет значительно повысить разрешающую способность при восстановлении изображения; 2) извлекает достаточно полную информацию из весьма ограниченных данных измерений (это свойство отчасти связано с первым); 3) обладает устойчивостью по отношению к шумам и неточностям измерений.

С теоретической точки зрения весьма важным было формальное доказательство того факта, что *метод максимума энтропии* (Maximum Entropy Method - MEM) является единственным методом, который не вносит новых ограничений на искомую оценку оригинала по сравнению с теми, которые обусловлены самими результатами измерений (Шор и Джонсон, 1980, 1983; Гэлл и Скиллинг, 1983; Ливерси и Скиллинг, 1984).

Все перечисленные факторы явились причиной широкого распространения на практике информационных методов восстановления изображения и вообще решения обратных задач. Эти методы применялись при анализе изображений, полученных космическими аппаратами, в ядерной энергетике, медицине, радиоастрономии и т.п. Что касается применений к равномерным временным рядам, то MEM позволил избавиться от боковых лепестков, характерных для классического Фурье-анализа, добиться уверенного выделения скрытых периодичностей в весьма коротких рядах и по крайней мере в несколько раз повысить разрешение (рис. 2.5.1, 2.5.2). К недостаткам MEM следует отнести трудность оценки амплитуды гармонической составляющей по виду MEM-спектра и, строго говоря, невозможность анализа неравномерных временных рядов.

Рассматриваемому кругу вопросов посвящено большое число исследований. Важнейшие понятия теории информации ясно описаны Шенноном в статьях 1948 и 1949 гг., вошедших в переведенный на русский язык сборник его работ (К.Шеннон, 1963), а в общедоступном изложении - в книге А.Яглома и И.Яглома (1973). Помимо работ, перечисленных в начале данного параграфа, необходимые сведения об анализе временных рядов с помощью MEM содержатся в работах Берга (1972), Лакосса (1971), Ульриха (1972 a,b), Ульриха и Бишопа (1975), Ульриха и Клейтона (1976), Фэлмена и Ульриха (1982), Кирка, Рас-

та и ван Винкля (1979), тематическом выпуске Proc. of IEEE (1982), книге Василенко и Тараторина (1986) и цитированной в них литературе. Астрономические приложения МЕМ обсуждаются в работах Эйблса (1974), Перси (1977), Фэлмена и Ульриха (1975, 1976).

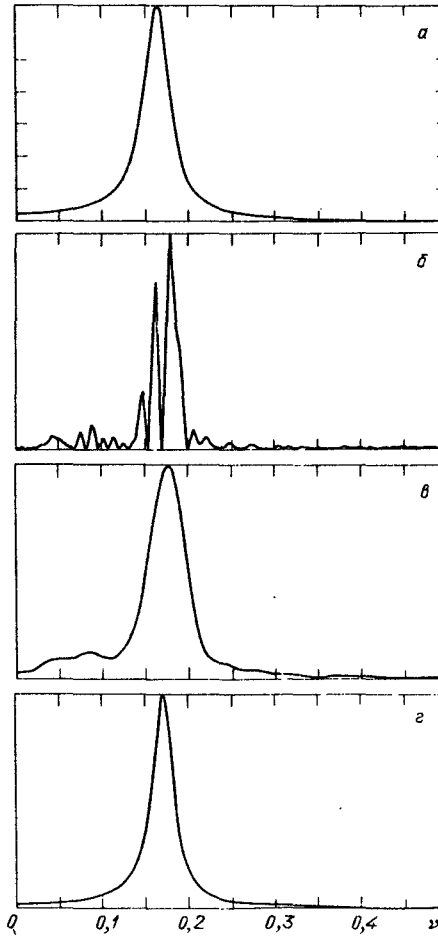


Рис. 2.5.1. а) Истинная спектральная плотность процесса, изображенного на рис. 2.6.1; б) периодограмма Шустера реализации, представленной на рис. 2.6.1; в) сглаженная с окном Бартлетта периодограмма (M = 25); г) МЕМ-спектр реализации рис. 2.6.1 при M = 2

Энтропия и информация. Дискретная случайная величина ξ полностью задана, если известны ее возможные значения $x_1, x_2, \ldots$ и соответствующие этим значениям вероятности их реализации $p_1, p_2, \ldots$ Энтропия $S(\xi)$ - детерминированная функция, характеризующая степень неопределенности, связанную со случайной величиной ξ: более «размытому» распределению вероятностей $\{p_k\}$ отвечает и большее значение энтропии. Определение $S(\xi)$ согласно К. Шеннону (1948), опирающееся на введенное Л. Больцманом (1872) соответствующее понятие статистической физики, имеет вид

$$S(\xi) = -\sum_{k=1}^{\infty} p_k \log p_k, \qquad (2.5.1)$$

где основанием логарифма чаще всего выбирают 2 (при этом S измеряется в *битах*). Детерминированной величине $\xi \equiv x_n$ соответствует распределение вероятностей $p_k = 0$ при $k \neq n$; $p_n = 1$ и энтропия $S = 0$. При конечном числе N возможных значений ξ максимальная энтропия отвечает равномерному распределению $p_k = 1/N$, $k = 1, 2, \ldots, N$; она равна $S_{max} = \log N$.

Приведем физический пример, который понадобится нам при рассмотрении теории счета фотонов в гл. 5. Известно (см. например, Лаудон, 1973, гл. 1), что число фотонов ξ в заданной моде теплового источника подчиняется вероятностному распределению

$$p_k = \frac{a^k}{(1+a)^{1+k}}, \quad k = 0, 1, 2, \ldots, \qquad (2.5.2)$$

которое называется *геометрическим распределением* (Феллер, 1966). Параметр $a \geqslant 0$ есть среднее число фотонов. Распределению (2) отвечает энтропия $S = \log [(1+a)^{1+a}/a^a]$. Представляет интерес то, что геометрическое распределение имеет наибольшую энтропию среди всех дискретных распределений $\{p_k\}$ с фиксированным средним значением $a = \sum k p_k$.

Обратим внимание, что дискретная случайная величина $\eta = f(\xi)$, являющаяся произвольной функцией ξ, имеет ту же энтропию, что и ξ.

Рассмотрим теперь непрерывную случайную величину ξ, принимающую значения в области $-\infty < x < \infty$ и характеризуюшуюся плотностью распределения вероятностей $p(x)$. Было бы естест-

венным попытаться обобщить определение энтропии (1) на случай непрерывной величины путем обычного предельного перехода, однако при этом возникает логарифмическая расходимость, имеющая очевидный смысл: для точного задания непрерывной величины необходима бесконечно большая информация (например, бесконечное количество десятичных знаков). Поскольку практически всегда интерес представляет не абсолютное значение энтропии, а разность энтропий двух случайных величин, целесообразно ввести так называемую *дифференциальную энтропию*

$$S(\xi) = -\int_{-\infty}^{\infty} p(x)\ln p(x)dx, \qquad (2.5.3)$$

где основанием логарифма принято число $e = 2,71828...$, а переменные нормированы так, что плотность $p(x)$ безразмерна.

Экспоненциально распределенная случайная величина с плотностью $p(x) = a^{-1}e^{-x/a}$ имеет среднее значение $\langle\xi\rangle = a$ и дифференциальную энтропию $S(\xi) = \ln(ae)$; это значение энтропии максимально на множестве всех распределений $p(x)$ с заданной средней величиной a. Аналогом этого результата из области статистической физики является распределение Больцмана систем по энергии $p(E) \propto e^{-E/\bar{E}}$ при заданной средней энергии $\bar{E}$. Для непрерывных распределений с фиксированным вторым моментом максимальной энтропией обладает нормальное распределение с дисперсией σ^2; его энтропия равна $S(\xi) = \ln(\sigma\sqrt{2\pi e})$. Вновь обращаясь к интерпретации этого результата на языке статистической физики, можно сказать, что максвелловское распределение молекул по скоростям обладает наибольшим статистическим весом при заданной температуре газа.

Таким образом, случайные величины с равномерной, экспоненциальной и нормальной плотностями распределения являются наиболее «неопределенными» при соответствующих ограничениях на совокупность значений или первые моменты.

Необходимо указать, что дифференциальная энтропия (3) не обладает некоторыми свойствами обычной энтропии. В частности, дифференциальная энтропия может принимать отрицательные значения; это свидетельствует лишь о том, что изучаемое

распределение вероятностей является более сосредоточенным по сравнению с некоторым референтным распределением. Кроме того, преобразование $\eta = f(\xi)$ непрерывной случайной величины ξ изменяет дифференциальную энтропию:

$$S(\eta) = S(\xi) + \int_{-\infty}^{\infty} \ln |f'(x)| p_{\xi}(x) dx. \qquad (2.5.4)$$

Ограничиваясь для определенности непрерывными распределениями, введем *совместную энтропию* произвольных случайных величин ξ и η соотношением

$$S(\xi,\eta) = -\iint_{-\infty}^{\infty} p(x,y) \ln p(x,y) dxdy, \qquad (2.5.5)$$

где $p(x,y)$ - плотность вероятности совместного распределения пары (ξ,η). Как известно, функция $p(x,y)$ представима в виде

$$p(x,y) = p_1(x)p(y|x) = p_2(y)p(x|y), \qquad (2.5.6)$$

где $p_1(x)$ и $p_2(y)$ - безусловные плотности распределения ξ и η, а $p(x|y)$ и $p(y|x)$ - соответствующие условные плотности при заданных η и ξ. Сопоставим функции $p(x|y)$ величину

$$S_c(\xi|\eta) = -\int_{-\infty}^{\infty} p(x|y) \ln p(x|y) dx, \qquad (2.5.7)$$

которая интерпретируется как условная энтропия ξ при заданной реализации $\eta = y$. Усредняя (7) по всем возможным значениям y, мы получим *среднюю условную энтропию* ξ при реализованной η:

$$S(\xi|\eta) = -\iint_{-\infty}^{\infty} p(x,y) \ln p(x|y) dxdy. \qquad (2.5.8)$$

Из соотношений (5), (6) и (8) нетрудно получить правило сложения энтропий:

$$S(\xi,\eta) = S(\eta) + S(\xi|\eta). \qquad (2.5.9)$$

Вообще говоря,

$$S(\xi|\eta) \leqslant S(\xi), \qquad (2.5.10)$$

и знак равенства здесь достигается при независимых ξ и η, когда $p(x|y) = p_1(x)$. Смысл неравенства (10) вполне понятен: знание результата измерения величины η сообщает опре-

деленную информацию и о распределении величины ξ, т.е. уменьшает степень неопределенности ξ. По этой причине уменьшение энтропии

$$I(\xi,\eta) \equiv S(\xi) - S(\xi|\eta) \qquad (2.5.11)$$

может рассматриваться в качестве меры информации о случайной величине ξ, содержащейся в опыте по реализации η. Следуя принятой терминологии, будем называть функцию $I(\xi,\eta)$ *информацией* о ξ, содержащейся в η. Естественно, что независимость величин ξ и η обусловливает равенства $S(\xi|\eta) = S(\xi)$, $I(\xi,\eta) = 0$, а правило сложения энтропий (9) принимает при этом вид

$$S(\xi,\eta) = S(\xi) + S(\eta). \qquad (2.5.12)$$

В общем случае, зависимы ξ и η или нет, информация $I(\xi,\eta)$ является симметричной функцией, т.е. $I(\eta,\xi) = I(\xi,\eta)$.

Рассмотрим в качестве примера случайные величины (ξ,η), подчиняющиеся двумерному нормальному распределению с коэффициентом корреляции ρ. Пусть σ_1^2 и σ_2^2 - дисперсии величин ξ и η соответственно. Выше мы отмечали, что безусловная энтропия $S(\xi) = \ln(\sigma_1\sqrt{2\pi e})$. Простые выкладки, которые мы не будем приводить, показывают, что средняя условная энтропия равна в данном случае $S(\xi|\eta) = \ln[\sigma_1\sqrt{2\pi e(1-\rho^2)}]$, так что

$$I(\xi,\eta) = -\frac{1}{2}\ln(1-\rho^2), \qquad |\rho| \leqslant 1. \qquad (2.5.13)$$

Как и следовало ожидать, для некоррелированных величин $I = 0$, а по мере усиления связи между ξ и η, т.е. при $|\rho|\to 1$, информация бесконечно возрастает.

В дальнейшем нам понадобится знание дифференциальной энтропии стационарного случайного процесса $\xi(t)$, и прежде всего той ее части, которая связана со спектральной плотностью процесса $g(\nu)$. К указанному понятию можно придти, рассматривая сначала среднюю энтропию N сечений процесса $\xi_1 = \xi(t_1),\ldots,\xi_N = \xi(t_N)$, имеющих совместную плотность вероятности $p(x_1,\ldots,x_N)$, т.е.

$$S_N[\xi(t)] = -\frac{1}{N}\int_N\ldots\int p(x_1,\ldots,x_N)\ln p(x_1,\ldots,x_N)\,dx_1\ldots dx_N, \qquad (2.5.14)$$

и устремляя затем N к бесконечности. Как и в одномерном случае (4), переход к новой системе случайных величин $\eta_k = f_k(\xi_1, \ldots, \xi_N)$, $k = 1, \ldots, N$, приводит к изменению дифференциальной энтропии

$$\Delta S_N = \frac{1}{N} \int_N \ldots \int \ln \left| J\left(\frac{f}{x}\right) \right| \; p(x_1, \ldots, x_N) dx_1 \ldots dx_N, \qquad (2.5.15)$$

где $J(f/x)$ - якобиан преобразования. В частности, при линейном преобразовании случайных величин с постоянными коэффициентами a_{kl} якобиан равен определителю N-го порядка $\det(a)$, составленному из коэффициентов преобразования, т.е. не зависит от переменных $x_1, \ldots, x_N$. Поэтому формула (15) принимает для линейного преобразования вид:

$$\Delta S_N = \frac{1}{N} \ln|\det(a)|. \qquad (2.5.16)$$

С целью найти ту часть энтропии, которая связана с конкретным видом спектральной плотности $g(\nu)$ случайного процесса, рассмотрим случайный спектр $F_\xi(\nu)$ процесса $\xi(t)$, определенный формулой (1.5.18). К функции $F_\xi(\nu)$, интерпретируемой как случайный процесс, зависящий от частоты ν, также применимы соотношения (14) - (16). В частности, при линейной фильтрации согласно (1.5.34) спектр $F_\eta(\nu)$ процесса $\eta(t)$ просто умножается на передаточную функцию $H(\nu)$ фильтра, т.е. $F_\eta(\nu) = H(\nu)F_\xi(\nu)$, где $H(\nu)$ определена формулой (1.5.40). Но при таком линейном преобразовании определитель, фигурирующий в (16), имеет диагональный вид, откуда следует $\det(a) = H(\nu_1)H(\nu_2)\cdot \ldots \cdot H(\nu_N)$ и

$$\Delta S_N = \frac{1}{N} \sum_{j=1}^{N} \ln|H(\nu_j)| = \sum_{j=1}^{N} \ln|H(\nu_j)|^2 \Delta\nu_j, \qquad (2.5.17)$$

где найквистовский интервал частот полагается разделенным на N равных промежутков $\Delta\nu_j = (2N)^{-1}$. Переходя в (17) к пределу при $N \to \infty$, мы получаем *формулу Шеннона* (1948)

$$\Delta S = \int_0^{\nu_c} \ln |H(\nu)|^2 d\nu, \qquad (2.5.18)$$

выражающую изменение энтропии при фильтрации случайного процесса.

Предположим теперь, что на вход линейного фильтра

поступает шум с постоянной спектральной плотностью. Тогда, согласно (1.5.41), на выходе фильтра сигнал имеет спектральную плотность $g(\nu) \propto |H(\nu)|^2$, а связанное с этим изменение энтропии с точностью до постоянного множителя представляется в виде:

$$\Delta S = \int_{-\nu_c}^{\nu_c} \ln g(\nu) d\nu. \qquad (2.5.19)$$

Эта формула является искомой. Хотя полное значение энтропии случайного процесса остается неизвестным, можно утверждать, что часть энтропии, обусловленная видом спектральной плотности, определяется выражением (19).

Подход Дж.Берга к проблеме оценивания спектральной плотности. С теоретической точки зрения задача нахождения спектральной плотности $g(\nu)$ бесконечного временного ряда $\{\xi_k\}$, $k = \ldots, -1, 0, 1, \ldots$ решается следующим образом (см. § 2.1 и § 2.3). Автоковариационная функция $\psi_n = \langle \xi_k \xi_{k+n} \rangle$, $n = \ldots, -1, 0, 1, \ldots$ связана с $g(\nu)$ соотношением

$$\psi_n = \int_{-\nu_c}^{\nu_c} e^{i \cdot 2\pi\nu\Delta t n} g(\nu) d\nu. \qquad (2.5.20)$$

Знание бесконечной совокупности $\{\psi_n\}$ позволяет найти точное решение уравнений (20) в виде:

$$g(\nu) = \Delta t \sum_{n=-\infty}^{\infty} e^{-i \cdot 2\pi\nu\Delta t n} \psi_n, \qquad (2.5.21)$$

где функция $g(\nu)$ положительно определена.

На опыте мы получаем лишь одну из возможных реализаций ряда $\{\xi_k\}$, которая к тому же имеет конечную длину. Обозначим наблюдаемую реализацию посредством $(y_0, y_1, \ldots, y_{N-1})$ и вычислим выборочную автоковариационную функцию

$$c_n = \frac{1}{N} \sum_{k=0}^{N-1-n} y_k y_{k+n}, \quad n = 0, \pm 1, \ldots, \pm(N-1), \qquad (2.5.22)$$

где по определению $c_{-n} = c_n$. Величины $\{c_n\}$ можно рассматривать как оценки первых N коэффициентов автоковариации $\{\psi_n\}$, однако ясно, что конечного набора $\{c_n\}$ недостаточно для строгого восстановления $g(\nu)$. Кроме того, нужно учитывать,

что оценка $g(\nu)$ неустойчива относительно небольших вариаций $\{c_n\}$, обусловленных самой случайной природой рассматриваемого процесса и ошибками измерений.

Традиционный подход к решению обсуждаемой задачи был кратко освещен в предыдущем параграфе. Этот подход основан на следующем доопределении автоковариационной функции:

$$\psi_n = \begin{cases} \lambda_n c_n , & |n| \leqslant M, \\ 0 , & |n| > M, \end{cases} \qquad (2.5.23)$$

т.е. наблюдаемые значения $\{c_n\}$ сглаживаются домножением на подходящие коэффициенты $\{\lambda_n\}$ вплоть до некоторого максимального радиуса корреляции $M \leqslant N-1$, а при $|n| > M$ значения ψ_n просто полагаются равными нулю. Однако при этом, во-первых, искажается имеющаяся информация о процессе (если λ_n не равны тождественно 1) и, во-вторых, делается конкретное предположение о недоступной информации - значениях ψ_n при $|n| > N-1$, причем нет никаких оснований считать, что выбор нулевых значений ψ_n оптимален (далее мы увидим, что такой выбор действительно нельзя считать хорошим).

Оба высказанных соображения очерчивают весьма существенные недостатки традиционного подхода к оцениванию спектральной плотности. Берг (1967) предложил новый, интуитивно ясный способ оценивания спектральной плотности, требующий, чтобы оценка $g(\nu)$ имела максимальную энтропию по сравнению с любой другой возможной оценкой спектральной плотности, согласующейся с результатами измерений.

С математической точки зрения задача сводится к нахождению такой функции $g(\nu) \geqslant 0$, которая доставляет максимум функционалу (19), т.е.

$$\int_{-\nu_c}^{\nu_c} \ln g(\nu)\,d\nu = \max \qquad (2.5.24)$$

при заданных условиях (*связях*)

$$\int_{-\nu_c}^{\nu_c} e^{i\cdot 2\pi\nu\Delta t n}\, g(\nu)\,d\nu = \psi_n, \quad |n| \leqslant M, \qquad (2.5.25)$$

предполагающих точное знание первых коэффициентов автоковариации.

Решение поставленной задачи проводится следующим образом

(см., в частности, превосходное изложение в работах Ульриха и Бишопа, 1975, и Кирка, Раста и ван Винкля, 1979). Заметим прежде всего, что из положительной определенности $g(\nu)$ следует аналогичное свойство и функции $1/g(\nu)$. Разложим $1/g(\nu)$ в ряд Фурье на промежутке $[-\nu_c, \nu_c]$ с неизвестными пока коэффициентами γ_n:

$$\frac{1}{g(\nu)} = \frac{1}{\Delta t} \sum_{n=-\infty}^{\infty} e^{-i\cdot 2\pi\nu\Delta t n} \gamma_n, \qquad (2.5.26)$$

где $\gamma_{-n} = \gamma_n$ и обращение имеет вид

$$\gamma_n = \Delta t \int_{-\nu_c}^{\nu_c} e^{i\cdot 2\pi\nu\Delta t n} \frac{\Delta t}{g(\nu)} d\nu. \qquad (2.5.27)$$

Найдем теперь максимум величины ΔS. Из (19), (21) и (27) следует:

$$\frac{\partial}{\partial\psi_n} \Delta S = \int_{-\nu_c}^{\nu_c} \frac{1}{g(\nu)} \frac{\partial g(\nu)}{\partial\psi_n} d\nu =$$

$$= \Delta t \int_{-\nu_c}^{\nu_c} e^{-i\cdot 2\pi\nu\Delta t n} \frac{1}{g(\nu)} d\nu = \frac{1}{\Delta t} \gamma_n. \qquad (2.5.28)$$

Таким образом, энтропия максимальна при выполнении условий

$$\gamma_n = 0, \quad |n| > M. \qquad (2.5.29)$$

Но тогда формула (26) принимает вид

$$g(\nu) = \frac{\Delta t}{\sum_{n=-M}^{M} e^{-i\cdot 2\pi\nu\Delta t n} \gamma_n}, \qquad (2.5.30)$$

а для нахождения $M+1$ коэффициентов $\gamma_0, \gamma_1, \ldots, \gamma_M$ служат условия (25). Можно показать, что эта последняя задача сводится к определению коэффициентов $a_1, a_2, \ldots, a_M$; P_M из системы линейных уравнений

$$\begin{aligned}
\psi_0 &= a_1\psi_1 + a_2\psi_2 + \ldots + a_M\psi_M + P_M, \\
\psi_1 &= a_1\psi_0 + a_2\psi_1 + \ldots + a_M\psi_{M-1}, \\
&\ldots \\
\psi_M &= a_1\psi_{M-1} + a_2\psi_{M-2} + \ldots + a_M\psi_0
\end{aligned} \qquad (2.5.31)$$

и последующему вычислению $g(\nu)$ по формуле

$$g(\nu) = \frac{P_M \Delta t}{\left| 1 - \sum_{m=1}^{M} e^{-i \cdot 2\pi\nu\Delta t m} a_m \right|^2}. \qquad (2.5.32)$$

При решении системы (31) в качестве $\{\psi_n\}$ можно было бы принять выборочные коэффициенты автоковариации $\{c_n\}$, вычисленные согласно (22). Однако выражение (22) предполагает, что $y_k \equiv 0$ при $k > N-1$, между тем как принцип максимальной энтропии требует отказа от предположений, касающихся недоступной информации. Кроме того, как показали Бокс и Дженкинс (1970), решение (31) весьма чувствительно к неточностям задания величин ψ_n, и в частности, ошибкам округления. Поэтому Берг (1968) разработал метод оценки параметров $a_1, \ldots, a_M$; P_M, не требующий предварительного вычисления коэффициентов автоковариации. Этот метод позволяет получить решение путем M-шагового рекурсивного алгоритма Левинсона (1947).

Наиболее удобным для вычисления спектральной плотности представляется алгоритм Андерсена (1974), который в случае временного ряда $(y_0, y_1, \ldots, y_{N-1})$ и заданного максимального радиуса корреляции M записывается следующим образом (см. Кирк и др., 1979).

1) На сетке $m = 1, 2, \ldots, M$; $j = 0, 1, 2, \ldots, N-1-m$ вычисляются вспомогательные коэффициенты $p_{j,m}$, $q_{j,m}$, $a_{m,m}$ и $P_{m,M}$. На первом шаге ($m = 1$)

$$p_{j,1} = y_j, \quad q_{j,1} = y_{j+1}, \quad j = 0, 1, \ldots, N-2,$$

$$a_{1,1} = \frac{2\sum_{j=0}^{N-2} y_j y_{j+1}}{\sum_{j=0}^{N-2} (y_j^2 + y_{j+1}^2)}, \qquad P_{1,M} = \frac{1-a_{1,1}^2}{N} \sum_{j=0}^{N-1} y_j^2. \qquad (2.5.33)$$

Далее при $m = 2, 3, \ldots, M$; $j = 0, 1, \ldots, N-1-m$ находим последовательно:

$$p_{j,m} = p_{j,m-1} - a_{m-1,m-1}\, q_{j,m-1},$$

$$q_{j,m} = q_{j+1,m-1} - a_{m-1,m-1}\, p_{j+1,m-1},$$

$$a_{m,m} = \frac{2\sum_{j=0}^{N-1-m} p_{j,m}\, q_{j,m}}{\sum_{j=0}^{N-1-m} (p_{j,m}^2 + q_{j,m}^2)}, \qquad (2.5.34)$$

$$P_{m,M} = P_{m-1,M}(1 - a_{m,m}^2).$$

2) На сетке $m = 2, 3, \ldots, M$; $k = 1, 2, \ldots, m-1$ вычисляются вспомогательные коэффициенты $a_{k,m}$ по формулам

$$a_{k,m} = a_{k,m-1} - a_{m,m}\, a_{m-k,m-1}, \qquad (2.5.35)$$

где $a_{m,m}$ известны по пункту 1.

3) Входящие в (32) постоянные $a_1, \ldots, a_M$ и P_M полагаются равными:

$$P_M = P_{M,M},$$
$$a_k = a_{k,M}, \quad k = 1, 2, \ldots, M. \qquad (2.5.36)$$

Вычисление $g(\nu)$ в (32) целесообразно выполнять при помощи алгоритма быстрого преобразования Фурье. Выбор значения M обсуждается ниже.

На рис. 2.5.2 в качестве примера представлены периодограмма Шустера отсчетов $\{y_k\}$ ряда A (см. приложение 7 и рис. 1.1.2) и МЕМ-спектры этого же ряда при $M = 10$, 22 и 35.

Выбор адекватного значения M составляет важную проблему, не получившую еще окончательного решения. При $M \ll N$ спектральное разрешение оказывается недостаточно высоким ввиду малой эффективной длины используемого отрезка ряда. С другой стороны, при значениях M порядка длины ряда N оценка спектра становится неустойчивой относительно появления ложных пиков (см., в частности, рис. 2.5.2, *г* и примеры в цитированных выше работах). Опыт показывает, что оптимальное значение M находится вблизи $N/2$; Берримен (1978) на основа-

нии численных экспериментов предложил формулу

$$M(N) \simeq \frac{2N}{\ln(2N)}, \qquad (2.5.37)$$

которая дает, например, M (50) $\simeq$ 22, M (100) $\simeq$ 38, M (500) $\simeq$ $\simeq$ 145. МЕМ-спектр ряда А на рис. 2.5.2 при M = 22, действительно, представляется оптимальным.

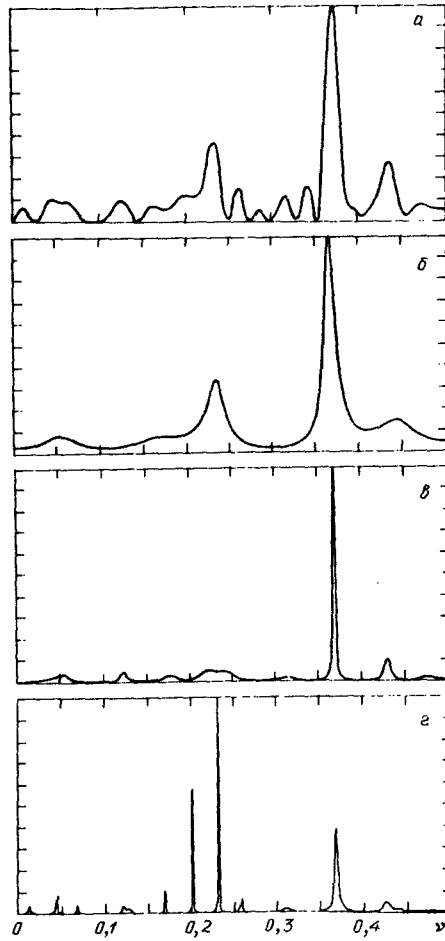


Рис. 2.5.2. Периодограмма Шустера $\{y_k\}$ ряда А (*а*) и МЕМ-спектры этого ряда при M = 10 (б), 22 (в) и 35 (*г*)

Аналитическое исследование проблемы выбора M предприняли Ульрих и Бишоп (1975), исходившие из требования минимальности введенной Акаике (1969, *а*, *б*; 1970) *окончательной*

ошибки предсказания (Final Prediction Error) для центрированного временного ряда:

$$FPE(M) = \frac{N + M + 1}{N - M - 1} P_M, \qquad (2.5.38)$$

Использование этого критерия предполагает последовательное повторение алгоритма (33) - (36) для всех M в интервале $[0, N-1]$ и окончательный выбор того значения, которое минимизирует $FPE(M)$. Обычно описываемая процедура дает разумные значения M, однако Кирк, Раст и ван Винкль (1978) приводят примеры, когда она ненадежна. Указанные авторы советуют руководствоваться при выборе M видом Фурье-спектра изучаемого временного ряда.

Одно из основных преимуществ метода максимальной энтропии связано со значительно более высоким разрешением, которое достигается в МЕМ-спектре по сравнению с периодограммой Шустера и тем более - сглаженным спектром. Именно высокое разрешение МЕМ-спектра позволяет, в частности, исследовать низкие частоты и выделять гармонические компоненты с периодом, превышающим длину исследуемого временного ряда. Реализуемое в методе максимальной энтропии «сверхразрешение», т.е. разделение компонентов с разницей частот $\delta\nu < 1/T$, не сразу нашло объяснение и производило на раннем этапе исследований несколько мистическое впечатление (см. рис. 3.8.4). Интерпретация явления «сверхразрешения» основана на следующих соображениях.

Прежде всего, параметры P_M и $a_1, \ldots, a_M$, определяющие согласно (32) форму МЕМ-спектра, допускают простое истолкование, связывающее метод максимальной энтропии с так называемыми *авторегрессионными моделями* временных рядов (ван ден Бос, 1971; § 2.6). Именно пусть

$$\hat{y}_k = \sum_{m=1}^{M} a_m y_{k-m} \qquad (2.5.39)$$

- линейная комбинация M предшествующих y_k значений временного ряда с некоторыми пока произвольными коэффициентами $a_1, \ldots, a_M$. Мы можем рассматривать $\varepsilon_k = y_k - \hat{y}_k$ как ошибку предсказания значения y_k и определить среднее качество

предсказания членов ряда как $P'_M = \langle \varepsilon_k^2 \rangle$. Можно показать, что требование минимальности P'_M приводит к системе уравнений (31) для нахождения $a_1, \ldots, a_M$, т.е. эти параметры представляют собой коэффициенты *оптимального линейного фильтра предсказания*, а само значение P_M есть минимальная среднеквадратическая ошибка линейного предсказания по M членам. Другими словами, при заданном M коэффициенты $a_1, \ldots, a_M$ дают наилучшее в смысле среднеквадратического отклонения представление временного ряда $\{y_k\}$ в виде

$$y_k = \sum_{m=1}^{M} a_m y_{k-m} + \varepsilon_k, \qquad (2.5.40)$$

где дисперсия шума ε_k равна P_M.

Традиционный подход к оценке спектра предполагает периодическое продолжение ряда вне границ отрезка времени длиной T, где действительно производились наблюдения. В противоположность этому метод максимума энтропии использует оптимальное приближение наблюдаемого отрезка ряда, которое допускает продолжение в обе стороны за границы интервала T. Из всех возможных продолжений, согласующихся с опытом, этот метод выбирает продолжение, обладающее наибольшим статистическим весом. Таким образом, эффективная длина временного ряда при построении МЕМ-спектра превосходит T, а соответствующее разрешение $\delta\nu < 1/T$.

Если процесс включает гармоническую составляющую с частотой ν_0 и амплитудой A, то, как показывает (1.6.8), среднее значение периодограммы Шустера на частоте ν_0 равно $A^2/4$, т.е. пропорционально энергии колебания. Эта пропорциональность не соблюдается для МЕМ-спектра, дающего оценку линейчатого спектра с более высоким разрешением. Кроме того, максимумы спектральной плотности, отвечающие гармоникам, обнаруживают «спонтанное расщепление» по мере уменьшения шага по частоте, зависящее от длины ряда и его начальной фазы (Фуже и др., 1976). Ясное обсуждение причины явления расшепления линий содержится в работе Джейнса (1982). Указанные обстоятельства следует учитывать при интерпретации наблюдательных данных и результатов численных экспериментов.

§ 2.6. Модели авторегрессии

Все рассмотренные нами до сих пор методы оценивания спектральной плотности процесса $g(\nu)$ на основании данных наблюдений не касались собственно механизма, обусловливающего конкретный вид $g(\nu)$ - задача заключалась в нахождении оценки, как можно более близкой (хотя бы при $N \to \infty$) к истинной спектральной плотности. Было выяснено, что периодограмма Шустера не может служить состоятельной оценкой $g(\nu)$; в качестве таковой следует использовать сглаженную периодограмму либо применить метод максимальной энтропии.

Предположим, что тем или иным способом «хорошая» оценка $g(\nu)$ для временного ряда длины N получена. Дальнейшая задача, непосредственно примыкающая к выяснению физических процессов, ответственных за наблюдаемую реализацию, состоит в *построении формальной модели* временного ряда. Например, в простейшем случае $g(\nu)$ может сравнительно мало меняться с частотой за исключением нескольких выделяющихся пиков, значимость которых доказана при помощи статистических критериев. Тогда вполне естественно принять в качестве модели процесса суперпозицию периодической функции и белого шума. Если с ростом N ширина спектральных линий и дисперсия оценки спектральной плотности убывают, у нас есть все основания приступить к физическому анализу процесса.

Рассмотрим, однако, несколько более сложную ситуацию, которая очень часто имеет место в астрономии. Временной ряд, изображенный на рис. 2.6.1, характеризуется явственно ощущаемой псевдопериодичностью, отдельные колебания прослеживаются в течение длительного времени, но амплитуда, частота и фаза этих колебаний испытывают хаотические вариации. В таких случаях исследователи часто говорят, что они «видят» периодичность, хотя в спектре мощности отсутствует узкая спектральная линия. Действительно, линия будет все более замываться по мере увеличения длины ряда, и его спектральная плотность примет вид, указанный на рис. 2.5.1, *а*. Казалось бы, возможности дальнейшего анализа такого процесса ограничены, но это не так: весьма эффективной является

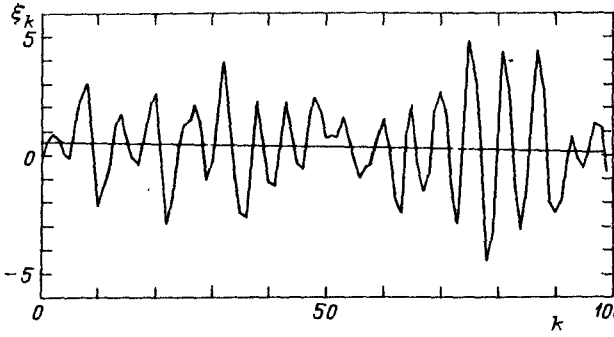


Рис. 2.6.1. Реализация авторегрессионного процесса AR(2) при $a_1 = 0,90$, $a_2 = -0,80$

конкретизация структуры процесса путем введения даже небольшого числа вспомогательных параметров. По этой причине такие модели называются *параметрическими моделями случайных процессов*. Так, отсчеты ξ_k временного ряда, показанного на рис. 2.6.1, были получены при помощи рекуррентного соотношения

$$\xi_k = a_1\xi_{k-1} + a_2\xi_{k-2} + \varepsilon_k, \qquad (2.6.1)$$

где a_1 и a_2 - постоянные (выбранные в данном случае равными 0,90 и -0,80 соответственно), ε_k - независимые и одинаково распределенные случайные величины (в примере - нормальные величины с нулевым средним и единичной дисперсией). Задание структуры процесса при помощи четырех параметров существенно экономнее описания всего множества значений спектральной плотности при каждом значении частоты.

Важно отметить, что соотношение (1) представляет собой уравнение в конечных разностях второго порядка. Подобно дифференциальным уравнениям второго порядка оно имеет колебательные решения (см. Хемминг, 1962, гл. 5), и именно это обстоятельство позволяет описывать псевдопериодические процессы.

Соотношение (1) задает *авторегрессионный процесс второго порядка*, или кратко AR(2). Его обобщением является процесс AR(p), который определяется линейным соотношением

$$\xi_k = a_1\xi_{k-1} + a_2\xi_{k-2} + \ldots + a_p\xi_{k-p} + \varepsilon_k, \qquad (2.6.2)$$

включающим p постоянных $a_1, \ldots, a_p$ и шум $\{\varepsilon_k\}$. *Авторегрессионные модели* были введены Г.Юлом (1927) в связи с рассмотрением данных о солнечной активности (см. § 2.7), и с тех пор стали основным инструментом при анализе геофизических и радиофизических временных рядов. Свое название авторегрессионные модели получили вследствие вида уравнения (2), связывающего обычно разные случайные переменные, но в данном случае - процесс $\xi(t_k)$ со значениями этого же процесса в предыдущие моменты времени. Процесс AR(0) есть просто шум $\{\varepsilon_k\}$.

Юл исходил из следующей наглядной модели. Пусть тяжелый маятник совершает колебания около положения равновесия. Если трение отсутствует, то дифференциальное уравнение второго порядка, описывающее движение маятника, не содержит члена, пропорционального скорости, и колебания продолжаются неограниченно долго. При движении, скажем, в воздухе необходимо ввести указанное слагаемое; движение маятника тогда имеет вид затухающих колебаний. Пусть далее маятник испытывает частые беспорядочные удары сравнительно небольших частиц, например, горошин. Существо дела состоит в том, что наблюдаемое движение маятника нельзя описать простой синусоидой постоянной или уменьшающейся амплитуды с наложенным на нее шумом. Действительное движение имеет вид *псевдоколебаний* с меняющимися частотой, фазой и амплитудой, причем достаточно высокая интенсивность случайных толчков компенсирует затухание и обеспечивает стационарность процесса. Высокочастотный шум сглаживается тенденцией маятника сохранять заданную моду колебания.

При помощи соотношения (2) описывается сохранение памяти о конечном числе прошлых состояний процесса ξ_k, однако с тем же правом можно предполагать сохранение памяти о бесчисленном количестве прошлых случайных воздействий $\{\varepsilon_j\}$, постепенно затухающих со временем. Соответствующее представление имеет вид:

$$\xi_k = \varepsilon_k + b_1\varepsilon_{k-1} + b_2\varepsilon_{k-2} + \ldots, \qquad (2.6.3)$$

где параметры $\{b_n\}$ характеризуют кривую затухания ($b_0 = 1$). Формально (3) можно получить из основного соотношения (2),

если вместо величины ξ_{k-1} подставить в правую его часть выражение

$$\xi_{k-1} = a_1\xi_{k-2} + a_2\xi_{k-3} + \ldots + a_p\xi_{k-p-1} + \varepsilon_{k-1}, \qquad (2.6.4)$$

следующее из того же определения AR(p), а затем неограниченно продолжить процесс последовательных подстановок. В частности, при подходящем выборе параметров a_1 и a_2 в уравнении (1) постоянные $\{b_n\}$ для процесса AR(2) имеют вид

$$b_n = \Lambda^n \frac{\sin[(n+1)\theta]}{\sin\theta}, \qquad n = 0, 1, 2, \ldots, \qquad (2.6.5)$$

где $\Lambda < 1$ и θ - новые постоянные, зависящие от a_1 и a_2 (см. соотношения (38) ниже). Мы видим, что кривая реакции на случайный импульс действительно описывается колебательной функцией времени с периодом $P_0 = 2\pi\Delta t/\theta$.

Частный случай общего представления (3), когда отличны от нуля только первые $q + 1$ коэффициентов $\{b_n\}$, т. е.

$$\xi_k = \varepsilon_k + b_1\varepsilon_{k-1} + \ldots + b_q\varepsilon_{k-q}, \qquad (2.6.6)$$

получил несколько неудачное название *процесса скользящего среднего* (Moving Average) порядка q, кратко MA(q).

Может показаться, что процессы AR(p) и MA(q) достаточно специфичны по своей структуре и поэтому не охватывают сколько-нибудь широкого класса реальных процессов. Это не так. Представление (3) описывает в терминах § 1.5 фильтрацию случайного процесса $\{\varepsilon_j\}$ фильтром с импульсной характеристикой

$$h(\tau) = \sum_{n=0}^{\infty} b_n\delta(\tau - n\Delta t), \qquad (2.6.7)$$

которой отвечает передаточная функция

$$H(\nu) = \sum_{n=0}^{\infty} b_n e^{-i\cdot 2\pi\nu\Delta t n}. \qquad (2.6.8)$$

Согласно (1.5.41), спектральные плотности входного $g_\varepsilon(\nu)$ и выходного $g_\xi(\nu)$ процессов (3) связаны соотношением

$$g_\xi(\nu) = |H(\nu)|^2 g_\varepsilon(\nu). \qquad (2.6.9)$$

Если входной процесс $\{\varepsilon_j\}$ - белый шум, то $g_\varepsilon(\nu) = \sigma_\varepsilon^2\Delta t =$

= const, и мы получаем

$$g_\xi(\nu) = \left| \sum_{n=0}^{\infty} e^{-i\cdot 2\pi\nu\Delta t n}\, b_n \right|^2 \sigma_\varepsilon^2 \Delta t. \qquad (2.6.10)$$

Ясно, что подбирая в (10) подходящую последовательность коэффициентов $\{b_n\}$, т.е. конструируя нужный фильтр, можно добиться удовлетворительного описания практически любого наблюдаемого спектра $g_\xi(\nu)$ (индекс ξ мы далее часто будем опускать).

Для процесса MA(q) сумма в (10) включает конечное число q слагаемых. Для процесса же AR(p) вычисление этой суммы требует, вообще говоря, знания бесконечного числа коэффициентов $\{b_n\}$ и потому не представляет практического интереса. Было бы желательным найти такое представление спектральной плотности $g(\nu)$ процесса AR(p), которое содержало бы только конечную совокупность коэффициентов $a_1,\ldots,a_p$, входящих в определение (2). Это выполняется следующим образом.

Обратим преобразование (1.5.18) и запишем спектр $F_\xi(\nu)$ процесса $\{\xi_k\}$ и спектр $F_\varepsilon(\nu)$ процесса $\{\varepsilon_k\}$ в виде

$$F_\xi(\nu) = \sum_{k=-\infty}^{\infty} e^{-i\cdot 2\pi\nu\Delta t k}\, \xi_k, \qquad (2.6.11)$$

$$F_\varepsilon(\nu) = \sum_{k=-\infty}^{\infty} e^{-i\cdot 2\pi\nu\Delta t k}\, \varepsilon_k.$$

Домножим уравнение (2) на $e^{-i\cdot 2\pi\nu\Delta t k}$ и суммируем по k от $-\infty$ до $+\infty$. Тогда получим

$$A(\nu) F_\xi(\nu) = F_\varepsilon(\nu), \qquad (2.6.12)$$

где

$$A(\nu) = 1 - a_1 e^{-i\cdot 2\pi\nu\Delta t} - a_2 e^{-i\cdot 2\pi\nu\Delta t\cdot 2} - \ldots - a_p e^{-i\cdot 2\pi\nu\Delta t\cdot p}. \qquad (2.6.13)$$

Произведем далее в (12) замену ν на ν', операцию комплексного сопряжения, умножим получившееся уравнение на (12) и усредним результат по ансамблю. Принимая во внимание (1.5.20), находим

$$g_\xi(\nu)\delta(\nu-\nu') A(\nu) A^*(\nu') = g_\varepsilon(\nu)\delta(\nu-\nu'). \qquad (2.6.14)$$

Искомое соотношение следует отсюда после интегрирования по ν' в бесконечных пределах:

$$g_\xi(\nu)\,|A(\nu)|^2 = g_\varepsilon(\nu). \qquad (2.6.15)$$

Мы далее повсюду будем считать $\{\varepsilon_j\}$ белым шумом с дисперсией σ_ε^2, так что спектральная плотность процесса AR(p) равна

$$g(\nu) = \frac{\sigma_\varepsilon^2 \Delta t}{\left|1 - \sum_{m=1}^{p} e^{-i\cdot 2\pi\nu\Delta t\cdot m}\, a_m\right|^2}. \qquad (2.6.16)$$

Полученное выражение совпадает с представлением (2.5.32) спектральной плотности в методе максимальной энтропии. Смысл этого результата обсуждался в предыдущем параграфе. Сейчас для нас важно то, что при надлежащем выборе порядка процесса авторегрессии p мы автоматически получим спектральную плотность, оптимальную с точки зрения метода максимальной энтропии.

Пока ограничимся этим в изложении общей теории процессов авторегрессии произвольного порядка и рассмотрим детально процессы AR(1) и AR(2), представляющие наибольший интерес для приложений. Нам понадобится для этих процессов: 1) совокупность коэффициентов $\{b_n\}$ в представлении (3); 2) автоковариационная функция процесса $\{\psi_n\}$, и в частности, его дисперсия $\psi_0 = \sigma_\xi^2$; 3) выражения коэффициентов корреляции $\{\rho_n\}$ через коэффициенты $\{a_k\}$ и обратные выражения $\{a_k\}$ через коэффициенты корреляции; 4) вид спектральной плотности $g(\nu)$.

Детальное рассмотрение авторегрессионных процессов содержится в монографиях Бокса и Дженкинса (1970), Андерсона (1971), статьях Ульриха и Бишопа (1975), Ульриха и Клейтона (1976), Скаргла (1981). Приложения процессов авторегрессии к астрономическим проблемам обсуждаются в последующих двух параграфах.

Модель AR(1) - марковский ряд. Понятие *марковского случайного процесса* оказалось весьма полезным при описании

многих физических явлений. Указанный процесс характеризуется тем свойством, что задание его состояния в некоторый момент времени определяет дальнейшее вероятностное поведение процесса безотносительно ко всему прошлому множеству состояний. Линейную дискретную модель марковского процесса можно представить в виде AR(1), т.е.

$$\xi_k = a_1\xi_{k-1} + \varepsilon_k, \tag{2.6.17}$$

где a_1 - постоянная и ε_k - белый шум с нулевым средним и дисперсией σ_ε^2:

$$\langle\varepsilon_k\rangle = 0, \quad \langle\varepsilon_k\varepsilon_l\rangle = \delta_{kl}\sigma_\varepsilon^2. \tag{2.6.18}$$

По предположению, величины $\{\xi_k\}$ образуют стационарный случайный процесс, так что $\langle\xi_k\rangle = 0$.

В данном случае процесс последовательных подстановок, приводящий к формуле (3), выполняется элементарно. Мы получаем в итоге:

$$\xi_k = \varepsilon_k + a_1\varepsilon_{k-1} + a_1^2\varepsilon_{k-2} + a_1^3\varepsilon_{k-3} + \ldots, \tag{2.6.19}$$

так что коэффициенты $\{b_n\}$ равны

$$b_n = a_1^n, \quad n = 0, 1, 2, \ldots \tag{2.6.20}$$

При

$$|a_1| < 1 \tag{2.6.21}$$

действие каждого случайного импульса ε_j затухает со временем по закону (20), и в каждый момент времени значение процесса ξ_k определяется суммой «хвостов», обусловленных всеми предыдущими импульсами. Полагая условие (21) выполненным, находим из (20):

$$|b_n| = e^{-n/n_1}, \quad n_1 = \frac{1}{-\ln|a_1|} > 0 \tag{2.6.22}$$

затухание происходит по экспоненциальному закону с характерным временем $\tau_1 = n_1\Delta t = \Delta t/(-\ln|a_1|)$.

Найдем автоковариационную функцию процесса $\{\xi_k\}$. Домножая (17) сначала на ξ_{k-n}, $n = 1, 2, \ldots$, затем на ξ_k и производя усреднение, получаем

$$\psi_n = \langle \xi_k \xi_{k-n} \rangle = a_1 \psi_{n-1}, \quad n = 1, 2, \ldots,$$

$$\psi_0 = \langle \xi_k^2 \rangle = a_1^2 \psi_0 + \sigma_\varepsilon^2, \tag{2.6.23}$$

откуда

$$\psi_n = a_1^n \psi_0, \quad n = 1, 2, \ldots,$$

$$\psi_0 = \sigma_\xi^2 = \frac{\sigma_\varepsilon^2}{1 - a_1^2}. \tag{2.6.24}$$

По определению, коэффициент корреляции ρ_n между состояниями процесса, разделенными интервалом времени $n\Delta t$, равен ψ_n/ψ_0, т.е. $\rho_n = a_1^n = b_n$, - величине, определенной соотношением (20). Коэффициент корреляции между соседними значениями процесса часто обозначается просто через ρ:

$$a_1 = \rho_1 \equiv \rho , \tag{2.6.25}$$

и мы приходим к совершенно ясной интерпретации параметра авторегрессии a_1.

Принимая во внимание (25), перепишем основные соотношения для процесса AR(1) в виде

$$\xi_k = \rho\, \xi_{k-1} + \varepsilon_k, \quad |\rho| < 1,$$

$$\rho_n = \rho^n, \qquad n = 0, 1, 2, \ldots, \tag{2.6.26}$$

$$\sigma_\xi^2 = \frac{\sigma_\varepsilon^2}{1 - \rho^2} .$$

При положительных значениях ρ автокорреляционная функция ρ_n экспоненциально убывает с ростом n, при $\rho < 0$ она на каждом следующем шаге меняет знак, сохраняя экспоненциальный закон убывания для $|\rho_n|$, как это показывает (22).

Нам осталось найти спектральную плотность $g(\nu)$ процесса AR(1). С учетом (25) имеем из общей формулы (16)

$$g(\nu) = \frac{\sigma_\varepsilon^2\, \Delta t}{1+\rho^2 - 2\rho \cos(2\pi\nu\Delta t)}, \quad |\nu| \leqslant \nu_c. \tag{2.6.27}$$

Поведение спектральной плотности марковского процесса при различных значениях коэффициента корреляции ρ иллюстрирует рис. 2.6.2.

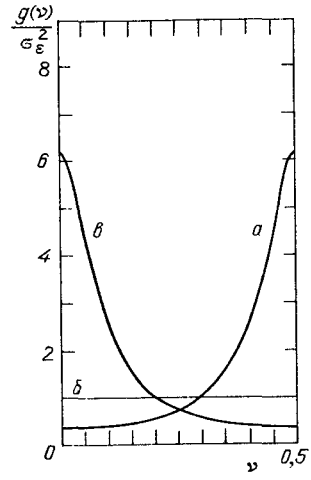


Рис. 2.6.2. Спектральная плотность процесса AR(1) при ρ= -0,6 (а), 0(б) и +0,6 (в)

Модель AR(2) - ряд Юла. Экспоненциальное затухание реакции со временем, свойственное марковскому ряду (17), обусловлено с формальной точки зрения тем, что разностное уравнение (17) имеет первый порядок. Если мы хотим описать явления типа псевдоколебаний, то должны ввести разностное уравнение второго порядка

$$\xi_k = a_1\xi_{k-1} + a_2\xi_{k-2} + \varepsilon_k, \qquad (2.6.28)$$

где ε_k по-прежнему представляет собой белый шум, описываемый соотношениями (18).

Для нахождения коэффициентов $\{b_n\}$ в представлении (3) домножим уравнение (28) последовательно на ε_k, ε_{k-1}, ε_{k-2} и т.д., а затем усредним результат. Поскольку из (3) следует

$$\langle\xi_k\varepsilon_{k-n}\rangle = b_n\sigma_\varepsilon^2, \quad n = 0, 1, 2, \ldots, \qquad (2.6.29)$$

получаем

$$b_0 = 1, \quad b_1 = a_1,$$
$$b_n = a_1 b_{n-1} + a_2 b_{n-2}, \quad n = 2, 3, \ldots \qquad (2.6.30)$$

Разностное уравнение (30) легко решается стандартными методами (Хемминг, 1962; Гельфонд, 1967). Именно, будем искать решение вида $b_n = \lambda^n$, где λ - неизвестная постоянная. Подставляя это решение в (30), получаем *характеристическое*

уравнение

$$\lambda^2 - a_1\lambda - a_2 = 0, \qquad (2.6.31)$$

корни которого суть

$$\lambda_1 = \frac{1}{2}\left(a_1 + \sqrt{a_1^2 + 4a_2}\right),$$
$$\lambda_2 = \frac{1}{2}\left(a_1 - \sqrt{a_1^2 + 4a_2}\right). \qquad (2.6.32)$$

Как и для дифференциальных уравнений, функции λ_1^n и λ_2^n являются частными решениями разностного уравнения (30), а его общее решение имеет вид

$$b_n = D_1\lambda_1^n + D_2\lambda_2^n, \qquad n = 0, 1, 2, \ldots, \qquad (2.6.33)$$

где D_1 и D_2 - произвольные постоянные. Учитывая равенства

$$a_1 = \lambda_1 + \lambda_2, \qquad a_2 = -\lambda_1\lambda_2, \qquad (2.6.34)$$

начальные условия (30) для нахождения D_1 и D_2 можно переписать в виде: $b_0 = 1$, $b_1 = \lambda_1 + \lambda_2$. Отсюда следует

$$D_1 = \frac{\lambda_1}{\lambda_1 - \lambda_2}, \qquad D_2 = -\frac{\lambda_2}{\lambda_1 - \lambda_2} \qquad (2.6.35)$$

и

$$b_n = \frac{\lambda_1^{n+1} - \lambda_2^{n+1}}{\lambda_1 - \lambda_2}, \qquad n = 0, 1, 2, \ldots \qquad (2.6.36)$$

Если корни характеристического уравнения действительны, то коэффициент b_n представляет собой суперпозицию двух показательных функций. Особого внимания с точки зрения физических приложений заслуживает случай комплексных сопряженных корней характеристического уравнения ($a_1^2 + 4a_2 < 0$). При этом можно принять

$$\lambda_1 = \Lambda e^{i\theta}, \qquad \lambda_2 = \Lambda e^{-i\theta},$$
$$|a_1|/2 < \Lambda < 1, \qquad 0 < \theta < \pi, \qquad (2.6.37)$$

где ограничения на Λ мы обсудим позднее. Постоянные авторегрессии a_1, a_2 связаны с Λ, θ соотношениями

$$\begin{cases} a_1 = 2\Lambda \cos\theta, \\ a_2 = -\Lambda^2, \end{cases} \quad \begin{cases} \Lambda = \sqrt{-a_2}, \\ \cos\theta = \dfrac{a_1}{2\sqrt{-a_2}}. \end{cases} \tag{2.6.38}$$

Подстановка (38) в (36) дает для случая комплексных корней выражение (5), физический смысл которого уже обсуждался во вводной части параграфа.

Наша следующая задача - найти автоковариационную функцию $\{\psi_\eta\}$ процесса AR(2). С этой целью домножим уравнение (28) последовательно на ξ_k, ξ_{k-1}, ξ_{k-2},..., а затем произведем усреднение. Учитывая, что $\psi_{-n} = \psi_n$, получаем

$$\begin{aligned} \psi_0 &= a_1\psi_1 + a_2\psi_2 + \sigma_\varepsilon^2, \\ \psi_1 &= a_1\psi_0 + a_2\psi_1, \\ \psi_2 &= a_1\psi_1 + a_2\psi_0, \\ \psi_n &= a_1\psi_{n-1} + a_2\psi_{n-2}, \quad n = 3, 4, \ldots \end{aligned} \tag{2.6.39}$$

Первое из этих уравнений дает

$$\psi_0 = \sigma_\xi^2 = \frac{\sigma_\varepsilon^2}{1 - a_1\rho_1 - a_2\rho_2}, \tag{2.6.40}$$

где коэффициенты корреляции $\rho_n \equiv \psi_n/\psi_0$, $n = 0, 1, \ldots$ Два следующих уравнения составляют для случая $p = 2$ систему

$$\begin{aligned} \rho_1 &= a_1 + a_2\rho_1, \\ \rho_2 &= a_1\rho_1 + a_2, \end{aligned} \tag{2.6.41}$$

найденную Юлом (1927) и Уокером (1931). Наконец, последнее разностное уравнение (39) можно записать в виде

$$\rho_n = a_1\rho_{n-1} + a_2\rho_{n-2}, \quad n = 1, 2, \ldots, \tag{2.6.42}$$

где мы приняли во внимание равенства $\rho_0 = 1$, $\rho_{-n} = \rho_n$.

Итак, коэффициенты корреляции $\{\rho_n\}$ определяются уравнением (42), начальными условиями которого могут служить заданные значения ρ_1 и ρ_2. Эти последние связаны с коэффициентами авторегрессии уравнениями Юла-Уокера (41), решение которых относительно a_1 и a_2 имеет вид

$$a_1 = \frac{\rho_1(1-\rho_2)}{1-\rho_1^2}, \qquad a_2 = \frac{\rho_2-\rho_1^2}{1-\rho_1^2}, \tag{2.6.43}$$

так что дисперсия процесса (40) равна

$$\sigma_\xi^2 = \frac{1-\rho_1^2}{(1-\rho_2)(1-2\rho_1^2+\rho_2)}\,\sigma_\varepsilon^2. \tag{2.6.44}$$

При теоретических расчетах или моделировании процесса заданными чаще являются не ρ_1, ρ_2, а коэффициенты a_1, a_2. Для этого случая решение уравнений (41) дает

$$\rho_1 = \frac{a_1}{1-a_2}, \qquad \rho_2 = a_2 + \frac{a_1^2}{1-a_2}, \tag{2.6.45}$$

и подстановка (45) в (40) определяет дисперсию процесса как функцию a_1 и a_2:

$$\sigma_\xi^2 = \frac{1-a_2}{1+a_2}\,\frac{\sigma_\varepsilon^2}{(1-a_2)^2-a_1^2}. \tag{2.6.46}$$

Прежде чем проводить дальнейшее изучение коэффициентов корреляции рассмотрим области изменения основных параметров ρ_1, ρ_2 и a_1, a_2, определяющих процесс AR(2). Из выражения (44) видно, что конечность и положительность дисперсии процесса требует, чтобы выполнялись условия:

$$-1 < \rho_1 < 1, \quad \rho_2 < 1, \quad \rho_2 > 2\rho_1^2 - 1. \tag{2.6.47}$$

Аналогичное требование в применении к (46) приводит к условиям

$$-1 < a_2 < 1, \quad a_1 + a_2 < 1, \quad a_2 - a_1 < 1\ . \tag{2.6.48}$$

Неравенства (47) и (48) называют *условиями стационарности процесса* AR(2); область (48) представлена на рис. 2.6.3. На этом же рисунке обозначена область $a_1^2 + 4a_2 < 0$, для которой корни характеристического уравнения (31) комплексны и, следовательно, реакция имеет вид затухающих колебаний. Как легко видеть, в указанной области параметр Λ заключен в пределах $|a_1|/2 < \Lambda < 1$.

Решение уравнения (42) при заданных ρ_1 и ρ_2 проводится совершенно так же, как это было выполнено при исследовании уравнения (30). Окончательные выражения для коэффициентов

корреляции и автоковариации имеют вид:

$$\rho_n = \frac{(1-\lambda_2^2)\lambda_1^{n+1} - (1-\lambda_1^2)\lambda_2^{n+1}}{(\lambda_1-\lambda_2)(1+\lambda_1\lambda_2)}, \quad n = 0, 1, 2, \ldots, \qquad (2.6.49)$$

$$\psi_n = \frac{\sigma_\varepsilon^2}{(\lambda_1-\lambda_2)(1-\lambda_1\lambda_2)} \left(\frac{\lambda_1^{n+1}}{1-\lambda_1^2} - \frac{\lambda_2^{n+1}}{1-\lambda_2^2} \right), \quad n=0, 1, 2, \ldots \qquad (2.6.50)$$

При $a_1^2 + 4a_2 > 0$ корни характеристического уравнения (31) действительны, а функции ρ_n и ψ_n, как ранее b_n, являются суперпозициями показательных функций. Если выполняется противоположное неравенство, то корни имеют вид (37), и мы получаем из (49)

$$\rho_n = \Lambda^n \frac{\sin(n\theta + \Phi)}{\sin \Phi}, \quad n = 0, 1, 2, \ldots, \qquad (2.6.51)$$

где вспомогательный угол Φ определяется соотношениями

$$\operatorname{tg} \Phi = \frac{1+\Lambda^2}{1-\Lambda^2} \operatorname{tg} \theta = \frac{1-a_2}{1+a_2} \frac{\sqrt{|a_1^2+4a_2|}}{a_1}, \quad -\frac{\pi}{2} < \Phi < \frac{\pi}{2}. \qquad (2.6.52)$$

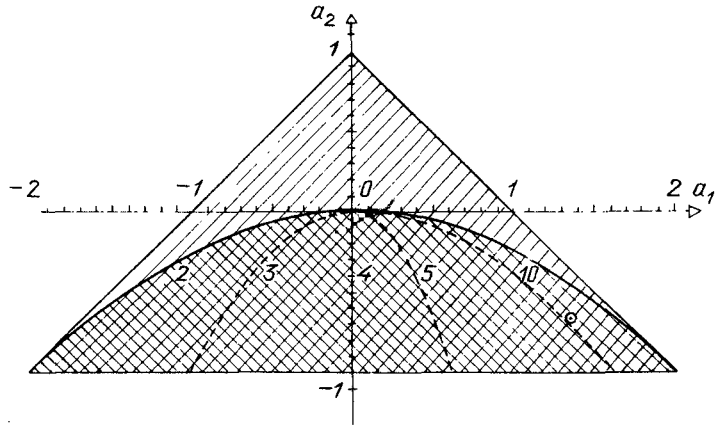


Рис. 2.6.3. Заштрихованная область представляет значения a_1 и a_2, при которых процесс AR(2) стационарен. Двойной штриховкой обозначена область, где корни характеристического уравнения комплексны. Цифрами отмечены линии постоянного периода P_0 = 2, 3, 4, 5 и 10; значком ⊙ – параметры авторегрессии пятнообразовательной деятельности Солнца, точкой – аналогичные значения для процесса на рис. 2.6.1

Таким образом, в области комплексных корней характеристического уравнения и реакция на случайный импульс, и коэффициенты корреляции имеют вид затухающих колебаний с периодом

$$P_0 = \frac{2\pi}{\theta}\Delta t = \frac{2\pi\Delta t}{\arccos\left(\dfrac{a_1}{2\sqrt{-a_2}}\right)}. \qquad (2.6.53)$$

Полагая $\Lambda^n = e^{-n/n_2}$, мы можем представить характерное время затухания τ_2 следующими выражениями:

$$\tau_2 \equiv n_2\Delta t = \frac{\Delta t}{-\ln\Lambda} = \frac{2\Delta t}{-\ln(-a_2)}. \qquad (2.6.54)$$

Напомним, что здесь предполагается $-1 < a_2 < -(a_1/2)^2$, так что $\tau_2 > \Delta t/\left(-\ln\dfrac{|a_1|}{2}\right)$.

Наконец, согласно (16), спектральная плотность процесса AR(2) равна

$$g(\nu) = \frac{\sigma_\varepsilon^2\Delta t}{1+a_1^2+a_2^2-2a_1(1-a_2)\cos(2\pi\nu\Delta t)-2a_2\cos(4\pi\nu\Delta t)}. \qquad (2.6.55)$$

При помощи этого выражения можно описать как монотонное уменьшение спектральной плотности с частотой, так и сравнительно узкие максимумы, характерные для квазипериодических колебаний (рис. 2.6.1).

Оценка параметров авторегрессии. До сих пор мы рассматривали аналитические свойства процессов AR(p) и не касались способов представления наблюдаемых реализаций временных рядов такими процессами. Между тем, на практике исследователь почти всегда имеет дело именно с этой последней проблемой, включающей в себя выбор порядка p и коэффициентов $a_1, \ldots, a_p$, σ_ε^2 модели AR(p), в определенном смысле наиболее близкой к данным наблюдений. Мы опишем здесь кратко два из применяющихся методов оценки параметров авторегрессии, отсылая за деталями к монографиям Бокса и Дженкинса (1970), Андерсона (1971) и цитированным выше статьям.

Первый из упомянутых методов восходит к оригинальным исследованиям Юла (1927) и Уокера (1931). Домножим уравнение (2) последовательно на ξ_k, $\xi_{k-1}, \ldots,$ ξ_{k-p} и усредним полу-

ченные соотношения. Принимая во внимание определение автоковариационной функции $\psi_n = \langle \xi_k \xi_{k-n} \rangle$, $n = 0, 1, \ldots$, находим систему уравнений

$$\psi_0 = a_1\psi_1 + a_2\psi_2 + \ldots + a_p\psi_p + \sigma_\varepsilon^2 , \qquad (2.6.56)$$

$$\begin{aligned} \psi_1 &= a_1\psi_0 + a_2\psi_1 + \ldots + a_p\psi_{p-1}, \\ \psi_2 &= a_1\psi_1 + a_2\psi_0 + \ldots + a_p\psi_{p-2}, \\ &\ldots \\ \psi_p &= a_1\psi_{p-1} + a_2\psi_{p-2} + \ldots + a_p\psi_0 , \end{aligned} \qquad (2.6.57)$$

фактически совпадающую с (2.5.31). Поскольку коэффициенты корреляции равны $\rho_n \equiv \psi_n/\psi_0$, имеем из (56)

$$\psi_0 = \sigma_\xi^2 = \frac{\sigma_\varepsilon^2}{1 - a_1\rho_1 - a_2\rho_2 - \ldots - a_p\rho_p} , \qquad (2.6.58)$$

а из (57) следует

$$\begin{aligned} \rho_1 &= a_1 + a_2\rho_1 + \ldots + a_p\rho_{p-1}, \\ \rho_2 &= a_1\rho_1 + a_2 + \ldots + a_p\rho_{p-2}, \\ &\ldots \\ \rho_p &= a_1\rho_{p-1} + a_2\rho_{p-2} + \ldots + a_p. \end{aligned} \qquad (2.6.59)$$

Как легко видеть, соотношения (56) - (59) представляют собой обобщение соотношений (39) - (41) на случай процесса произвольного порядка p. Соответственно система (59) называется *системой уравнений Юла-Уокера* для процесса AR(p).

Если теоретические значения ρ_n в (59) заменить выборочными значениями

$$\hat{\rho}_n = \frac{c_n}{c_0} , \quad c_n = \frac{1}{N}\sum_{k=0}^{N-1-n} y_k y_{k+n}, \qquad n = 0, 1, \ldots, p, \qquad (2.6.60)$$

а затем решить эту систему уравнений относительно $a_1, \ldots a_p$, то мы придем к так называемым *оценкам Юла-Уокера* параметров авторегрессии (оценка σ_ε^2 находится из (58) после замены ψ_0 на c_0). Недостатки этой процедуры уже обсуждались нами в предыдущем параграфе.

В настоящее время более употребительным является метод Берга (1968), основанный на глубокой связи между авторегрессионными моделями, теорией оптимального прогнозирования и методом максимальной энтропии (см. § 2.5). При этом сначала оценивается порядок модели либо при помощи вычисления

окончательной ошибки предсказания Акаике (2.5.38), либо на основании других соображений. Затем, если это не было сделано ранее при расчете FPE, находятся коэффициенты авторегрессии $a_1, \ldots, a_M$ и остаточная дисперсия шума $P_M = \sigma_\varepsilon^2$ при помощи алгоритма (2.5.33) - (2.5.36). Эти параметры полностью определяют подобранную авторегрессионную модель. При необходимости найти кривую реакции системы на импульсное воздействие, т.е. совокупность коэффициентов $\{b_n\}$ в представлении (3), можно последовательно использовать соотношения

$$b_0 = 1,$$

$$b_1 = a_1,$$

$$b_2 = a_1 b_1 + a_2 b_0,$$

$$b_3 = a_1 b_2 + a_2 b_1 + a_3 b_0, \qquad (2.6.61)$$

$$\ldots$$

$$b_p = a_1 b_{p-1} + a_2 b_{p-2} + \ldots + a_p b_0,$$

$$b_n = a_1 b_{n-1} + a_2 b_{n-2} + \ldots + a_p b_{n-p}, \quad n = p+1, \ldots, \qquad (2.6.62)$$

получающиеся при домножении (2) на ε_k, ε_{k-1}, ... и усреднении результатов с учетом (29).

Таким образом, сначала определяются коэффициенты AR(p)-представления и только затем - коэффициенты MA(q)-представления. Преимущества этого подхода обусловлены наличием процедуры выбора порядка авторегрессионного процесса и тем обстоятельством, что нахождение коэффициентов $\{a_m\}$ сводится к линейной задаче, в то время как прямое определение $\{b_n\}$ требует итераций нелинейных уравнений (Фэлмен и Ульрих, 1975; Бокс и Дженкинс, 1972).

Заметим, наконец, что процесс «гармонический сигнал плюс шум», нередко выбираемый в качестве теста при расчетах МЕМ-спектра и авторегрессионного приближения, невозможно описать чистой моделью авторегрессии (Ульрих и Клейтон, 1976), и этот факт следует учитывать при анализе контрольных примеров.

§ 2.7. Вариации солнечной активности

В 1843 г. любитель астрономии С.Швабе обнаружил, что число пятен на Солнце изменяется с течением времени приблизительно периодически с длиной цикла около 10 лет. Как выяснилось в дальнейшем, вариации числа пятен отражают общие колебания солнечной активности. В качестве текущей характеристики пятнообразовательной деятельности Солнца наблюдатели обычно используют *числа Вольфа*, которые для каждого дня определяются как сумма числа отдельных пятен и умноженного на 10 числа их групп. Именно ряд среднегодовых чисел Вольфа послужил Шустеру (1906) и Юлу (1927) для практической проверки введенных ими важнейших понятий - периодограммы и модели авторегрессии.

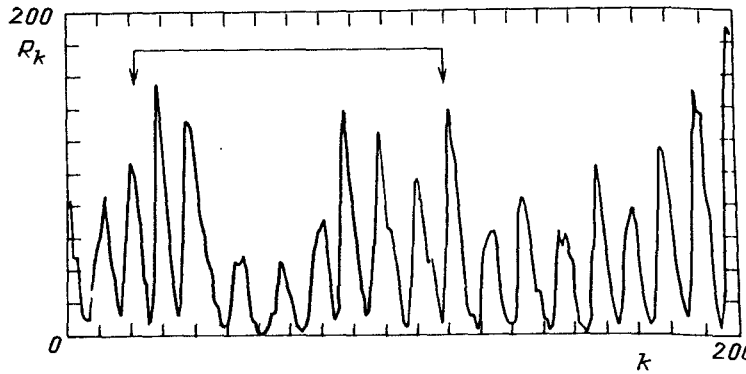


Рис. 2.7.1. Ряд W чисел Вольфа с 1749 по 1958 гг. (N= 210); стрелками выделен ряд E (N = 100) за период 1770 - 1869 гг.

В данном параграфе мы рассмотрим ряд W среднегодовых чисел Вольфа $\{R_k\}$ за период 1749-1958 гг. (всего 210 значений, см. табл. П.7.2 и рис. 2.7.1), а также его часть - ряд E, включающий указанные числа для периода с 1770 по 1869 гг. (100 значений). Одна из причин выделения ряда E связана с тем, что этот ряд детально изучался Боксом и Дженкинсом (1970) при помощи методов, предшествующих методу максимальной энтропии (МЕМ), и сопоставление полученных результатов представляет несомненный интерес. Другая причина заключается в желании выяснить устойчивость результатов при

расширении данных наблюдений.

Излишне говорить, что исследованию ряда чисел Вольфа посвящена огромная литература. Мы не предполагаем здесь обсуждать эти работы, поскольку наша цель состоит не столько в изучении собственно активности Солнца, сколько в иллюстрации практического применения основных методов.

Ряд E. Целесообразно рассмотреть сначала более короткий ряд *E*. Угловой коэффициент линейного тренда $\alpha_1 = -0,08 \pm 0,13$ нельзя считать в данном случае значимо отличающимся от нуля. Периодограмма Шустера центрированного ряда *E* представлена на рис. 2.7.2, *а*. Наиболее выделяющийся пик $P_N(\nu)$ соответствует частоте 0,086 цикл/год, т.е. периоду около

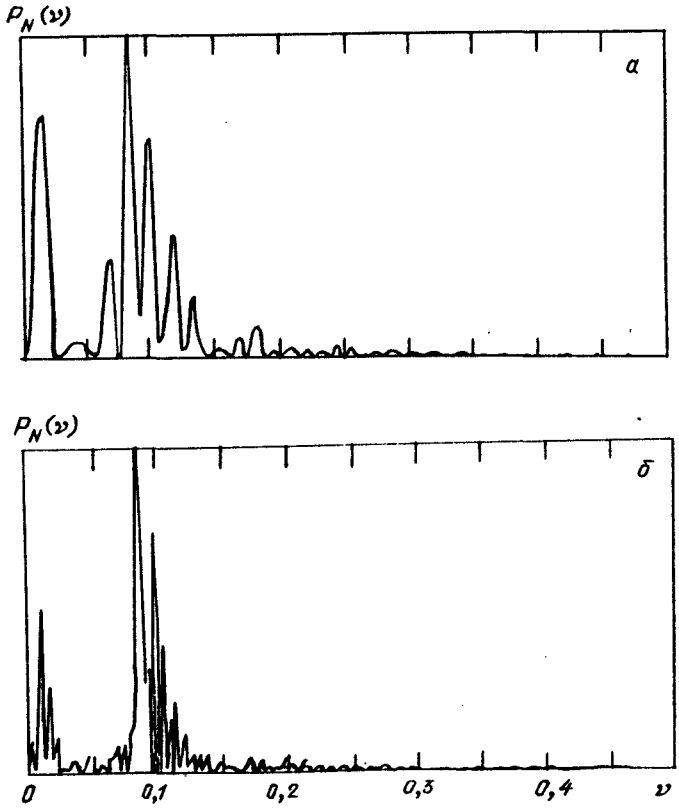


Рис. 2.7.2. Периодограммы Шустера рядов *E* (а) и *W* (б)

11,6 лет. Вблизи указанной частоты на периодограмме заметны пики несколько меньшей амплитуды, свидетельствующие, однако, не в пользу реальности набора близких периодов, а скорее в пользу размытости основного максимума. Это говорит о недостаточно четкой выраженности периода чисел Вольфа. В самом деле, рис. 2.7.1 показывает, что расстояние между экстремумами ряда изменяется в довольно широких пределах, отдельные циклы подвержены как бы сбоям частоты и амплитуды. На периодограмме выделяется и низкочастотный максимум, отвечающий периоду около 72 лет. Возможность существования долгой периодичности (P ~ 70 - 90 лет) впервые была отмечена Глейсбергом (см. Вязаницын и др., 1964).

Если вспомнить приведенные в предыдущем параграфе соображения, связанные с физическим смыслом основных уравнений авторегрессионной модели, то введение ее при интерпретации колебаний солнечной активности представляется совершенно естественным. Поскольку вычисление МЕМ-спектра автоматически приводит к одной из моделей авторегрессии, будем опираться при выборе порядка авторегрессии p (или M в обозначениях § 2.5) на требование минимальности окончательной ошибки предсказания Акаике (2.5.38). Параметр P_M был вычислен для ряда значений M при помощи алгоритма (2.5.33) - (2.5.36), а затем по формуле (2.5.38) определялась функция FPE(M). Ее значения представлены на рис. 2.7.3. Мы видим, что при переходе от $M = 1$ к $M = 2$ величина FPE(M) резко уменьшается, а при дальнейшем увеличении M медленно возрастает с небольшими флуктуациями. Ввиду этого можно полагать, что модель AR(2) в первом приближении правильно отражает основные особенности ряда чисел Вольфа. Конечно, включение моделей более высокого порядка позволит уточнить описание процесса (Крэддок, 1967). Мы ограничимся здесь рассмотрением модели AR(2).

В ходе описанных выше расчетов были вычислены и значения параметров a_1 и a_2, определяющих в соответствии с (2.6.28) процесс AR(2). Эти значения, а также дисперсия шума $P_M = \sigma_\varepsilon^2$ указаны в табл. 2.7.1. Любопытно сравнить их с аналогичными значениями, найденными из системы уравнений Юла-Уокера (2.6.41). Сначала необходимо найти по формулам (2.6.60) вы-

борочные коэффициенты корреляции; они оказались равными $\hat{\rho}_1 = 0,806$, $\hat{\rho}_2 = 0,428$. Подставляя эти значения в (2.6.43), получаем оценки Юла-Уокера (YW) параметров $\hat{a}_1 = 1,32$ и $\hat{a}_2 = -0,63$. Поскольку дисперсия ряда E равна $\sigma_R^2 = 1396,9$, из (2.6.44) можно найти оценку остаточной дисперсии шума $\sigma_\varepsilon^2 \simeq 294$ (см. табл. 2.7.1). Оценки YW коэффициентов авторегрессии оказались удаленными в пределах их естественных флуктуаций от оценок, найденных методом максимальной энтропии. Следует, однако, обратить внимание на то, что дисперсия шума для MEM-оценок значимо меньше, чем для оценок Юла-Уокера. Это свидетельствует о высокой эффективности метода максимальной энтропии.

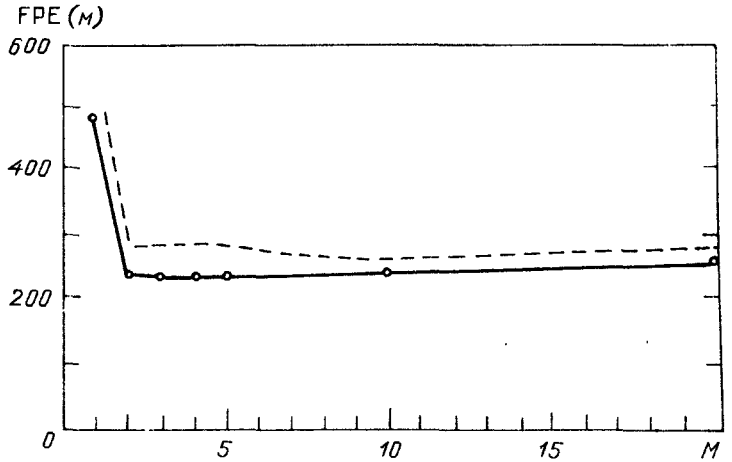


Рис. 2.7.3. Окончательная ошибка предсказания в зависимости от порядка модели авторегрессии для рядов E (сплошная линия) и W (штриховая)

Заметим в этой связи, что сложный итеративный метод наименьших квадратов, описанный Боксом и Дженкинсом (1970), приводит к оценкам параметров авторегрессии, весьма близким к MEM-оценкам (см. табл. 2.7.1).

Принимая значения параметров авторегрессии ряда E, найденные методом максимальной энтропии, нетрудно вычислить по формуле (2.6.55) и MEM-спектр изучаемого процесса. Он представлен на рис. 2.7.4, *а*. Спектральная плотность достигает максимума при частоте ~0,09 цикл/год (период ~11,1 лет).

Более точное описание периодических свойств процесса мы получим, если воспользуемся формулами (2.6.53). В данном случае находим $\theta = 33,82^{\circ}$ и $P_0 = 10,64$ года. Обратим внимание на значительную спектральную мощность в области малых частот. Это говорит о том, что для рассматриваемого процесса характерны также медленные колебания сравнительно небольшой амплитуды, не имеющие четко выраженного периода. С этой точки зрения не исключено, что долгопериодическая зависимость с длиной цикла ~80 лет, о которой упоминалось выше, является просто частным свойством одной из реализаций.

Таблица 2.7.1

Ряд	Метод	N	a_1	a_2	σ_ε^2
	YW	100	1,32±0,08	-0,63±0,08	294
E	LS	100	1,42±0,07	-0,73±0,07	228
(1770-1869 гг.)	MEM	100	1,40	-0,71	229
Юла					
(1749-1924 гг.)	YW	176	1,34	-0,66	237
W					
(1749-1958 гг.)	MEM	210	1,37	-0,68	272

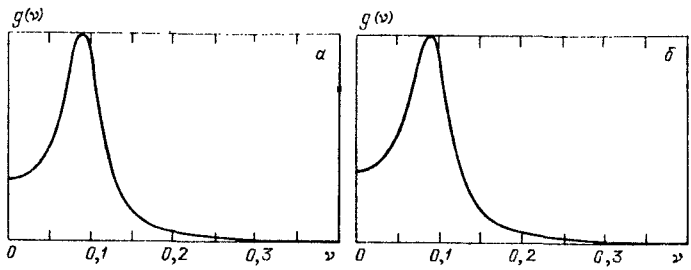


Рис. 2.7.4. MEM-спектры рядов E (а) и W (б)

Сам Юл изучал ряд чисел Вольфа с 1749 г. по 1924 г. ($N = 176$). Он получил выборочные коэффициенты корреляции $\hat{\rho}_1 = 0.811$, $\hat{\rho}_2 = 0,434$ и оценки параметров авторегрессии, указанные в табл. 2.7.1. При этом $\theta = 33,96^{\circ}$, $P_0 = 10,60$ лет, т.е. оценка периода солнечной активности практически совпадает с найденной выше.

Ряд W. Как и для ряда E, отличие углового коэффициента

линейного тренда ряда W от нуля нельзя считать значимым (α_1 = 0,050±0,044). Среднее значение $\bar{R}$ = 48,5±2,7.

Если авторегрессионная модель второго порядка правильно описывает внутреннюю структуру ряда чисел Вольфа, то более чем двукратное увеличение продолжительности ряда не должно существенным образом сказаться на основных характеристиках модели, приведенных в предыдущем пункте. Так оно и оказывается в действительности.

Зависимость FPE(*M*) для ряда W показана штриховой линией на рис. 2.7.3. Как и ранее, мы ограничимся рассмотрением модели AR(2). Соответствующие параметры, вычисленные при помощи метода максимальной энтропии, даны в табл. 2.7.1. Как видно, они мало отличаются от найденных для ряда *E*. Вычисления по формуле (2.6.53) дают значения θ = 33,83° и P_0 = = 10,64 года.

Рисунки 2.7.2, *б* и 2.7.4, *б* показывают, что общие черты распределения спектральной плотности сохранились при существенном увеличении длины ряда.

Наконец, выбирая оценки для ряда W как более надежные, мы опишем ряд чисел Вольфа $\{R_k\}$ при помощи модели (2.6.28):

$$R_k = 1,37R_{k-1} - 0,68R_{k-2} + 15,0 + 16,49u_k, \qquad (2.7.1)$$

где принято $\xi_k = R_k - \bar{R}$, $\varepsilon_k = \sigma_\varepsilon u_k$ и гауссовский белый шум u_k имеет нулевое среднее и единичную дисперсию. Более точное описание дается авторегрессионными моделями высших порядков.

§ 2.8. Анализ кривой блеска квазара 3C 273

Основным структурным образованием наблюдаемой части Вселенной являются *галактики* - системы, включающие в себя ~10^9 - 10^{12} звезд и сравнительно небольшую массу газа, объединенных взаимным тяготением. Обычно поверхностная яркость галактик возрастает к центру в соответствии с повышением звездной плотности, однако существуют галактики, в центре которых наблюдается дополнительный пик светимости малых угловых размеров (как говорят, *звездообразный*). Светимость центрального образования (*ядра*) составляет для близких

звездных систем величину порядка нескольких десятых долей светимости звездного компонента галактик. Первые галактики с яркими звездообразными ядрами были обнаружены в начале нашего века (Фэт, 1908; Слайфер, 1917), в качестве особого класса объектов они были выделены К. Сейфертом (1943), и получили название *сейфертовских галактик*. Большая ширина эмиссионных линий в спектрах ярких галактических ядер свидетельствует о скоростях движения газа порядка 10^3-10^4 км/с.

Галактики Сейферта достаточно многочисленны (они составляют около 1/100 всех ярких галактик), и потому встречаются даже среди близких звездных систем. В начале 60-х годов были обнаружены аналогичные объекты, светимость ядер в которых может на несколько порядков превосходить оптическую светимость звездного компонента (Шмидт, 1963; Сэндидж, 1965). Такие галактики назвали *квазарами* (Quasi Stellar Radio Sources - Quasars). Поскольку лишь малая часть галактических ядер находится в состоянии столь высокой активности, средняя пространственная плотность квазаров невелика, а соответствующее значение среднего расстояния до ближайших ярких квазаров составляет сотни мегапарсеков (Мпк = $3,086 \cdot 10^{24}$см). Именно вследствие удаленности квазаров первые наблюдения не выявили их оптической структуры, и в частности, наличия звездной составляющей. По этой же причине в спектрах квазаров отчетливо проявляется общее космологическое расширение Вселенной, приводящее к смещению линий в длинноволновую сторону спектра на тысячи ангстрем ($\Delta\lambda/\lambda_0 \leqslant 4,5$). Дальнейшие исследования выявили галактики, занимающие промежуточное в отношении светимости ядер положение между сейфертовскими галактиками и яркими квазарами.

Природа процессов, обусловливающих светимость ядер галактик $\sim 10^{45}$- 10^{48} эрг/с, не вполне ясна. Наиболее вероятным механизмом считается выделение энергии при аккреции газа *черной дырой* с массой $\sim 10^8 M_\odot$ (Солпитер, 1964; Зельдович, 1964; Линден-Белл, 1969).

Характерной особенностью активных ядер галактик является переменность потока излучения от них. Минимальное время пе-

ременности составляет несколько суток в оптическом диапазоне спектра и достигает нескольких часов - в рентгеновском. Амплитуда изменений оптического потока обычно не превосходит десятков процентов, но при возможности проследить фотометрическую историю в течение десятков лет отмечены вариации потока в несколько раз. Наиболее полные ряды наблюдений имеются для первого из обнаруженных квазаров - 3С 273, фотометрические измерения которого в 1887-1980 гг. сведены в одну систему (Смит и Хоффлейт, 1963; Энжион и Смит, 1985). Эти данные неоднократно служили предметом тщательного анализа, поскольку изучение кривой блеска позволяет высказать достаточно определенные суждения о природе источника энергии в активных ядрах галактик. В самом деле, доказательство, скажем, реальности квазипериодических вариаций потока с характерным временем ~10 лет, на которые указывали многие авторы (Смит и Хоффлейт, 1963; Смит, 1965; Озерной и Чертопруд, 1966; Кункель, 1967), послужило бы не только весомым аргументом в пользу единого излучающего тела, но и дало указания на характер протекающих физических процессов. Разумеется, допустить наличие квазипериодического компонента в изменениях блеска 3С 273 можно лишь при условии, что кривую блеска нельзя представить как результат наложения случайных флуктуаций потока. В пользу этой второй точки зрения свидетельствуют результаты анализа, выполненного впервые Филдом (1964), а затем более детально Мэнвеллом и Саймоном (1966, 1968), Террелом и Олсеном (1970), Фэлменом и Ульрихом (1975, 1976). При этом весьма важным для физической интерпретации представляется определение средней частоты независимых вспышек и их средней кривой блеска.

В настоящем параграфе мы рассмотрим возможность представления кривой блеска 3С 273 в виде суммы независимых импульсов определенной формы, случайно распределенных во времени. В радиофизике такой случайный процесс принято называть *дробовым шумом* (см., например, Рытов, 1976). В соответствии с цитированными выше работами будет показано, что имеющиеся данные наблюдений 3С 273 не противоречат модели дробового шума, и, таким образом, для принятия гипотезы о наличии периодического компонента блеска нет оснований.

Дробовой шум. Предположим, что случайный процесс $\xi(t)$ в каждый данный момент времени представляет собой сумму остаточных реакций («хвостов»), обусловленных бесконечным числом предыдущих импульсов и передних фронтов последующих импульсов:

$$\xi(t) = \sum_{k=1}^{\infty} a_k s\left(\frac{t-t_k}{\theta_k}\right), \qquad (2.8.1)$$

где a_k - случайная амплитуда импульса, достигшего максимума в момент t_k, θ_k - статистически не зависящая от a_k ширина импульса, а детерминированная функция $s(x)$ характеризует форму импульсов. Предположим, что амплитуды и длительности различных импульсов некоррелированы между собой, и введем обозначения

$$\langle a_k a_l \rangle = \begin{cases} \overline{a^2}, & k = l, \\ (\bar{a})^2, & k \neq l, \end{cases} \qquad \langle \theta_k \theta_l \rangle = \begin{cases} \overline{\theta^2}, & k = l, \\ (\bar{\theta})^2, & k \neq l. \end{cases} \qquad (2.8.2)$$

Пусть далее плотность распределения θ_k есть $u_\theta(\tau)$ и моменты $\{t_k\}$ образуют пуассоновскую последовательность, так что вероятность наступления j событий в течение промежутка времени Δt равна

$$\Pr\,[\mathfrak{N}(\Delta t) = j] = e^{-\bar{n}\Delta t}\,\frac{(\bar{n}\Delta t)^j}{j!}, \quad j = 0,\ 1,\ 2, \ldots, \qquad (2.8.3)$$

где $\bar{n}$ - средняя интенсивность возникновения импульсов. Наконец, пусть

$$S(z) = \int_{-\infty}^{\infty} e^{-i\cdot 2\pi z x}\, s(x)\,dx \qquad (2.8.4)$$

- трансформанта Фурье кривой блеска отдельного импульса. Приведем без вывода основные соотношения (так называемые *формулы Кемпбелла*, 1909, *a*, *b*), дающие среднее значение $\langle \xi(t) \rangle$, дисперсию $D[\xi(t)]$ и спектральную плотность процесса дробового шума:

$$\langle \xi(t) \rangle = \bar{n}\,\bar{a}\,\bar{\theta} \int_{-\infty}^{\infty} s(x)\,dx,$$

$$D[\xi(t)] = \bar{n}\,\overline{a^2}\,\bar{\theta} \int_{-\infty}^{\infty} s^2(x)\,dx, \qquad (2.8.5)$$

$$g(\nu) = \bar{n}\,\overline{a^2}\int_0^\infty |S(\nu\tau)|^2 u_\theta(\tau)\tau^2 d\tau, \quad 0 < \nu < \infty.$$

Вывод формул Кемпбелла и соответствующее обсуждение можно найти в монографии Рытова (1976). В качестве первого приближения часто рассматривается случай, когда параметры $\{a_k\}$ и $\{\theta_k\}$ фиксированы, т.е. флуктуациям подвержено лишь распределение импульсов во времени, а форма каждого из них определяется детерминированной функцией $a\mathfrak{s}[(t-t_k)/\theta] \equiv p(t-t_k)$. Тогда процесс дробового шума можно представить в виде

$$\xi(t) = \sum_k p(t-t_k), \tag{2.8.6}$$

а его характеристики следуют из (4) и (5):

$$\langle\xi(t)\rangle = \bar{n}\int_{-\infty}^{\infty} p(t)dt, \tag{2.8.7}$$

$$D[\xi(t)] = \bar{n}\int_{-\infty}^{\infty} p^2(t)dt, \tag{2.8.8}$$

$$g(\nu) = \bar{n}\left|\int_{-\infty}^{\infty} e^{-i\cdot 2\pi\nu t} p(t)dt\right|^2, \quad 0 < \nu < \infty. \tag{2.8.9}$$

Обратим внимание на то обстоятельство, что вид спектральной плотности дробового шума $g(\nu)$ определяется, согласно (9), лишь формой отдельного импульса $p(t)$. Этот результат становится очевидным, если учесть, что процесс (6) можно представить в виде свертки импульсного белого шума и функции $p(t)$; переход к спектральному представлению отвечает умножению постоянной спектральной плотности шума и спектральной плотности импульса.

При анализе данных наблюдений сравнительно просто находятся первые моменты $\langle\xi(t)\rangle$ и $D[\xi(t)]$. Если удается определить и форму импульса $p(t)$, то из (7) и (8) следует оценка средней частоты вспышек:

$$\bar{n} = \frac{\langle\xi(t)\rangle^2 \int p^2(t)dt}{D[\xi(t)]\left[\int p(t)dt\right]^2}. \tag{2.8.10}$$

Формула (10) применима и тогда, когда функция $p(t)$ первоначально найдена лишь с точностью до некоторой нормирующей

постоянной k, т.е. $p(t) = k^{-1}b(t)$; для нахождения k можно обратиться к уравнению (7).

Заметим, что если в (5) отвлечься от формы импульсов и дисперсии их амплитуд, то относительная вариация дробового шума определяется по порядку величины единственным параметром $N_0 \equiv \bar{n}\,\theta$, а именно:

$$\frac{\sigma_\xi}{\langle\xi\rangle} \sim \frac{1}{\sqrt{N_0}} \quad . \qquad (2.8.11)$$

Параметр N_0 - среднее число импульсов, возникших в течение средней длительности одного импульса, характеризует степень перекрытия импульсов в суммарном процессе. Можно сказать, что в данный момент времени значение процесса формируется суммой вкладов приблизительно N_0 отдельных импульсов. Так, для 3С 273 с его среднеквадратичной вариацией светимости ~ 14 % следует ожидать $N_0 \sim 50$, и мы получим ниже близкое значение на основании более тонкого анализа данных.

Запишем в качестве примера основные соотношения для случая импульсов экспоненциальной формы, т.е. $p(t) = 0$ при $t < 0$, $p(t) = ae^{-t/\theta}$ при $t \geqslant 0$. Мы имеем из (7) и (9)

$$\langle\xi\rangle = aN_0, \quad \frac{\sigma_\xi}{\langle\xi\rangle} = \frac{1}{\sqrt{2N_0}} ,$$

$$g(\nu) = \frac{\bar{n}a^2\theta^2}{(1+2\pi\nu\theta)^2} , \quad \nu > 0. \qquad (2.8.12)$$

Вид спектральной плотности при экспоненциальных вспышках иллюстрируется рис. 2.8.2, *в*.

Дискретной версией модели дробового шума может служить процесс скользящего среднего MA(q) с подобранными надлежащим образом характеристиками (см., например, Скаргл, 1981). Если Δt - временной шаг дискретизации, то случайное число событий пуассоновского процесса, происшедших на интервале Δt, можно интерпретировать как случайную амплитуду импульса дискретного процесса MA(q). Соотношение (3) дает тогда квантованное с единичным шагом вероятностное распределение амплитуд импульсов, форма которых $b(n)$ соответствует дискретному представлению функции $p(t)$ в (6).

В § 2.6 мы выяснили, что процесс скользящего среднего MA(q) и авторегрессионный процесс AR(p) тесно связаны между собой. Вместе с тем последний значительно проще сопоставить данным эксперимента, и потому при рассмотрении вариаций блеска 3C 273 естественно искать подходящую модель авторегрессии. В дальнейшем возможно восстановление исходного пуассоновского процесса (innovation), порождающего наблюдаемый временной ряд (Фэлмен и Ульрих, 1976).

Данные наблюдений. Начиная с 1963 г., когда 3C 273 был идентифицирован в качестве первого квазара, проводятся регулярные наблюдения вариаций его блеска при помощи техники фотоэлектрических измерений. За более ранний период 1955-1963 гг. также имеются, хотя и более редкие, фотоэлектрические оценки. Наконец, за период с конца прошлого века до начала 50-х годов имеются лишь фотографические измерения, уступающие по точности фотоэлектрическим. Смит и Хоффлейт (1963), а затем Кункель (1967) свели все имеющиеся данные с 1887 г. по 1967 г. в однородную систему, причем в последней работе путем линейной интерполяции звездных величин в пробелах был образован равномерный временной ряд, включающий 292 средние за 100-дневные отрезки времени оценки блеска 3C 273. Именно этот ряд явился в дальнейшем предметом всестороннего анализа (см. ссылки в начале данного параграфа). Сейчас ряд наблюдений 3C 273 в однородной системе доведен до 1980 г. (Энжион и Смит, 1985), т.е. охватывает почти столетие, однако, имея в виду методическую сторону, мы будем рассматривать преимущественно ряд Кункеля (табл. П.7.3).

Измерения в шкале звездных величин ряда Кункеля были переведены в линейные оценки, пропорциональные потоку излучения, согласно формуле $F = 10^{-0,4(m-m_0)}$ при $m_0 = 16,25$. Полученный равномерный ряд $\{F_k\}$ представлен на рис. 2.8.1, где указан также линейный тренд

$$F_k = \underset{\pm 0,37}{26,98} - \underset{\pm 0,0022}{0,0175}\, k\,, \qquad k = 0,\ 1,\ \ldots,\ 291. \tag{2.8.13}$$

Заметим, что k означает здесь число 100-дневных интервалов, прошедших с момента JD = 2 410 480, промежутки времени измеряются в единицах Δt = 100 сут, а частоты - в единицах

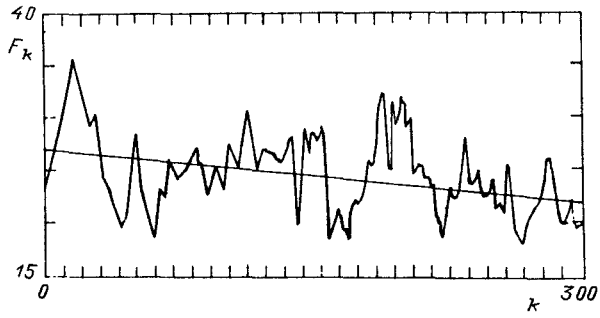


Рис. 2.8.1. Кривая блеска квазара 3C 273 в 1887 - 1967 гг. По оси абсцисс указано количество 100-дневных промежутков времени, прошедших с JD 2410480; по оси ординат - средние за эти интервалы значения потока излучения

цикл/100 сут = 3,6525 цикл/год. Из (11) нетрудно найти, что тренд составляет -26 % за столетие; существенная его часть обусловлена несколькими первыми годами, когда наблюдения были недостаточно часты и надежны. Если исключить первые 25 точек ряда (1887-1894 гг.), то величина тренда уменьшится до -16 % за столетие, что вполне совместимо с моделью случайных флуктуаций (Террел и Олсен, 1970).

Как показывают расчеты, дальнейшие результаты слабо зависят от того, используются ли данные, освобожденные от линейного тренда, или только центрированные значения потоков. Поэтому ниже мы рассматриваем данные последнего вида.

Оценка спектральной плотности. В качестве такой оценки для частот $\nu > 0$ естественно сначала принять периодограмму Шустера (рис. 2.8.2, *а*). Как и следовало ожидать, периодограмма имеет весьма изрезанный вид. Отдельные максимумы как-будто указывают на возможную периодичность ($P \sim 27$ лет, 16 лет и пр.), однако реальность дискретной составляющей спектра вызывает серьезные сомнения. В частности, расположение точек на фазовой диаграмме при возможных значениях периода имеет хаотический вид. Ясно, что необходима состоятельная оценка спектральной плотности, и сглаживание при помощи окна Бартлетта (2.4.25) действительно приводит к чисто непрерывному спектру, по-видимому, монотонно убывающему с ростом частоты (рис. 2.8.2, *б*).

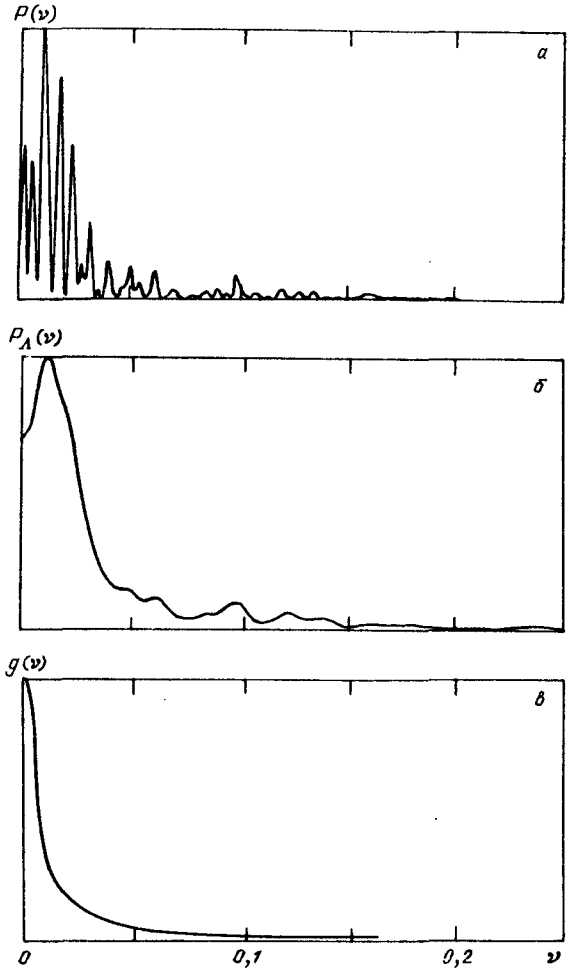


Рис. 2.8.2. а) Периодограмма кривой блеска 3С 273; б) та же периодограмма, сглаженная при помощи окна Бартлетта (M = 100); в) спектральная плотность пуассоновского импульсного процесса с экспоненциальной формой импульса

Наиболее подробное исследование вариаций светимости 3С 273, основанное на анализе спектральной плотности, предприняли Террел и Олсен (1970). Эти авторы пришли к выводу, что данные наблюдений удовлетворительно описываются моделью дробового шума, предполагающей наложение экспоненциальных вспышек с частотой $\bar{n} \simeq (15\pm5)$ год$^{-1}$ и средней продолжительностью $\theta \simeq (3,2\pm1)$ года. Еще ранее близкие значения ука-

занных параметров получил Филд (1964). Эффективное число налагающихся вспышек $N_0 = \bar{n}\theta \simeq 48 \pm 10$, а вид спектральной плотности определяется последним из соотношений (12). Соответствующая функция представлена на рис. 2.8.2, *в* с точностью до нормирующего множителя. Качественное согласие ожидаемой в области положительных частот для дробового шума и наблюдаемой спектральных плотностей вариаций блеска 3С 273 дает основания для более тщательного анализа модели.

Авторегрессионный анализ. Изменения светимости 3С 273 (рис. 2.8.1) и, в особенности, характер распределения спектральной плотности этого процесса (рис. 2.8.2) весьма сходны с таковыми для изучавшихся в § 2.6 процессов авторегрессии достаточно низкого порядка. Соответствующий анализ ряда Кункеля был выполнен Фэлменом и Ульрихом (1975). Результаты, близкие к полученным в этой работе, следуют и из несколько иной процедуры авторегрессионного анализа, описываемой ниже.

Прежде всего необходимо определить порядок p подходящей модели авторегрессии. Поскольку, как уже указывалось, метод максимальной энтропии (MEM) дает оценку спектральной плотности, совпадающую со спектральной плотностью одной из моделей авторегрессии, это было сделано следующим образом. При помощи алгоритма (2.5.33) - (2.5.36) для различных значений M были вычислены параметры P_M и $a_1, \ldots, a_M$, определяющие MEM-спектр. Затем по формуле (2.5.38) для каждого M определялась окончательная ошибка предсказания $\mathrm{FPE}(M)$. Минимум $\mathrm{FPE}(M)$ достигается при $M = 3$, что и отвечает согласно критерию Акаике оптимальному порядку авторегрессионной модели. Подставляя вычисленные значения коэффициентов $a_1 - a_3$ в определение (2.6.2) процесса AR(3), получаем:

$$\xi_k = 1,554\xi_{k-1} - 0,849\xi_{k-2} + 0,236\xi_{k-3} + \varepsilon_k, \qquad (2.8.14)$$

где дисперсия шума $\{\varepsilon_k\}$ равна $P_M = \sigma_\varepsilon^2 = 0,9565$.

Мы располагаем теперь значениями всех параметров, необходимых для нахождения спектральной плотности процесса по формулам (2.5.32) или (2.6.16). Эта плотность (оптимальная с точки зрения метода максимальной энтропии) представлена

на рис. 2.8.3. Функция $g(\nu)$ плавно спадает по мере увеличения частоты, не обнаруживая каких-либо выдающихся пиков, которые могли бы указывать на периодичность процесса (см. обсуждение в следующем пункте).

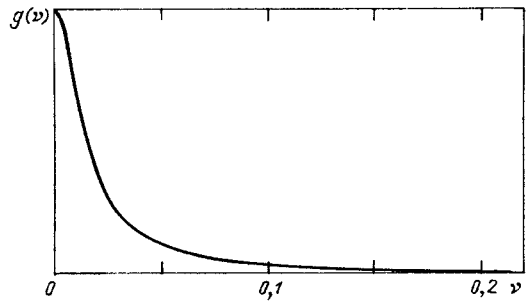


Рис. 2.8.3. MEM-спектр вариаций блеска 3C 273

Значительный интерес представляет определение формы отдельной вспышки. В рамках процессов AR(p) речь идет о нахождении коэффициентов $\{b_n\}$ в MA-представлении (2.6.3). Это может быть сделано при помощи формул (2.6.61) и (2.6.62). Результаты представлены на рис. 2.8.4, где, напомним, отсчеты по оси абсцисс означают число 100-дневных промежутков времени, прошедших с момента начала вспышки.

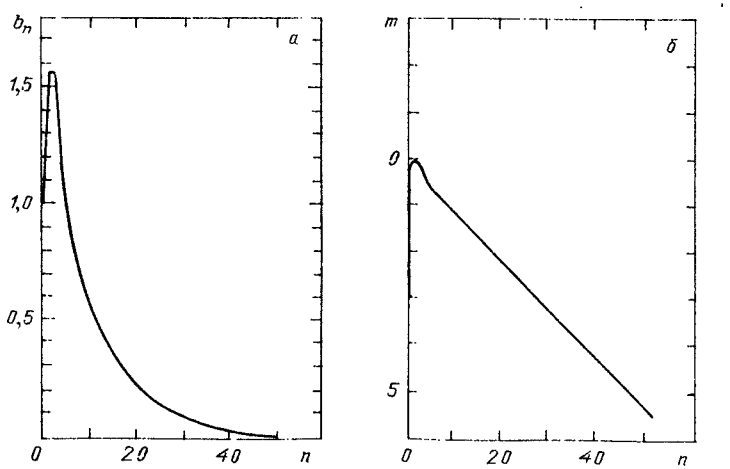


Рис. 2.8.4. Форма отдельной вспышки для 3C 273. а) Коэффициенты MA-представления, б) то же в шкале звездных величин

Наконец, знание кривой блеска отдельной вспышки позволяет оценить по формуле (10) и среднюю частоту их следования $\bar{n}$. Расчеты дают для рассматриваемого случая $\langle\xi\rangle = 24,44$, $D(\xi) = (3,73)^2$, $\sum_0^\infty b_n = 17,04$, $\sum_0^\infty n b_n = 165,65$, $\sum_0^\infty b_n^2 = 14,52$, и мы имеем

$$\bar{n} \simeq 2,2 \text{ 1/100 сут} \simeq 8 \text{ 1/год},$$
$$\bar{\theta} \simeq 9,7 \cdot 100 \text{ сут} \simeq 2,7 \text{ года}, \qquad (2.8.15)$$
$$N_0 \simeq 21.$$

Отличие приведенной оценки эффективного числа налагающихся вспышек N_0 от найденной в работе Террела и Олсена (1970) объясняется более точным учетом формы вспышек при авторегрессионной декомпозиции.

Обсуждение результатов. Итак, вариации светимости 3С 273 аппроксимированы простой моделью, описываемой уравнением (14). При этом саму наблюдавшуюся кривую блеска квазара следует трактовать как одну из возможных реализаций процесса AR(3) с параметрами (14). В этой связи интересно вычислить другие реализации процесса (14) и сопоставить их общий вид с кривой блеска квазара; практика свидетельствует, что простое визуальное изучение процесса дает удивительно много информации о его внутренней структуре. Для расчетов необходимо задать три начальных значения процесса, однако конкретный выбор их в известной степени произволен, так как процесс быстро «забывает» исходное состояние, если оно не слишком сильно отличается от среднего.

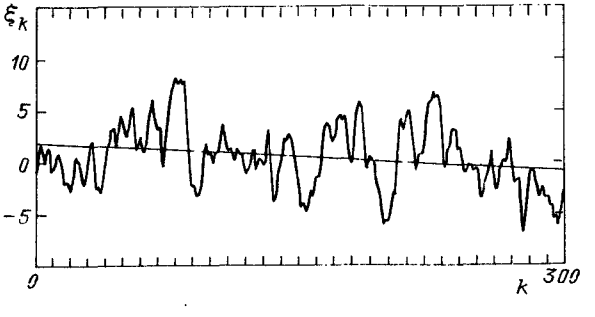


Рис. 2.8.5. Типичная реализация авторегрессионного процесса 3-го порядка, заданного формулой (14).

Одна из типичных реализаций процесса, заданного формулой (14), длины $N = 300$ представлена на рис. 2.8.5. Сопоставление его с рис. 2.8.1 показывает, что модель AR(3) правильно отражает все основные свойства действительного процесса вариаций блеска 3C 273. Заметим, что и тренд численной реализации случайно оказался близким к тренду ряда Кункеля. Выборочные авторегрессионные параметры отрезка временного ряда на рис. 2.8.5, а именно $a_1 = 1,522$, $a_2 = -0,797$, $a_3 = -0,184$, $\sigma_\varepsilon^2 = 0,914$ не сильно отличаются от исходных значений (14).

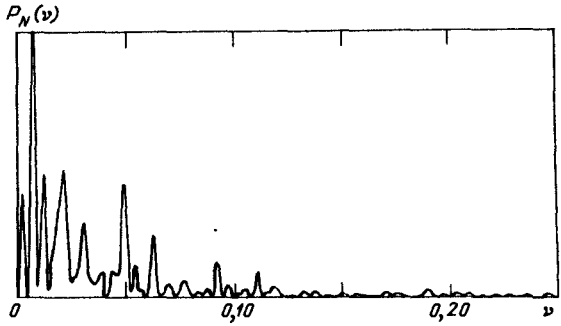


Рис. 2.8.6. Периодограмма Шустера реализации, показанной на рис. 2.8.5

Ввиду долгой дискуссии о реальности квазипериодических компонентов в кривой блеска 3C 273 представляет несомненный интерес вычисление периодограммы искусственного ряда AR(3), представленного на рис. 2.8.5. Она показана на рис. 2.8.6. Мы видим значительные случайные флуктуации периодограммы того же характера, что и реально наблюдавшиеся для 3C 273 (рис. 2.8.2, *а*). Аналогичное явление обнаружили Фэлмен и Ульрих (1975) в ходе численного моделирования дробового шума, заведомо не содержащего периодических компонентов. При сглаживании периодограммы все «квазипериодические детали» постепенно исчезают, что независимо свидетельствует о случайности их природы. Перечисленные факты дают основания сделать заключение о том, что максимумы периодограммы вариаций блеска 3C 273, которые можно было бы рассматривать как указание на существование квазипериодических колебаний, в действительности порождаются, скорее всего, случайными флук-

туациями, неизбежными для метода Шустера (см. § 2.3 - 2.5).

Таким образом, модель дробового шума со средней частотой вспышек порядка восьми в год и средней кривой блеска отдельной вспышки, изображенной на рис. 2.8.4, представляется вполне удовлетворительной при интерпретации изменений светимости 3С 273. Кривая блеска вспышки характеризуется крутым начальным ростом и весьма протяженным экспоненциальным спадом $b(t) \simeq 1{,}74e^{-t/9{,}77\Delta t}$, $t \geqslant 10\Delta t$, где Δt = = 100 сут. В шкале звездных величин второму этапу соответствует прямолинейный участок на рис. 2.8.4, *б* со скоростью убывания блеска 1^m/900 сут. Фэлмен и Ульрих (1975) получили близкую оценку скорости падения блеска 1^m/800 сут. Указанные авторы отмечают, что найденная средняя форма вспышки весьма напоминает оптические кривые блеска сверхновых звезд типа I, однако последние характеризуются приблизительно на порядок более быстрой временной шкалой и на три порядка меньшей освобожденной энергией (расстояние 3С 273 приблизительно равно 630 Мпк, и из рис. 2.8.4 нетрудно получить оценку $\Delta E \sim 10^{53}$ эрг).

Проведенный анализ как-будто исключает интерпретацию Филда (1964) переменности блеска квазаров в рамках модели, предполагающей суперпозицию вспышек обычных сверхновых, однако при взрывах сверхмассивных звезд (10 - 100 $M_\odot$) может освобождаться энергия вплоть до $\sim 10^{54}$ эрг (Колгейт, 1967; Колгейт и др., 1975). Все же представляется более вероятным инициирование случайных вспышек единым массивным образованием. В этой связи следует обратить внимание на модель Хиллса (1975), связывающую эволюцию центральных черных дыр в ядрах галактик с процессом приливного разрушения звезд. Подробное обсуждение относящихся сюда вопросов выходит за рамки данной книги (см. Хиллс, 1978; Рис, 1978; Теребиж, 1976, 1981).

В заключение параграфа подчеркнем, что использование усредненных по 100-дневным интервалам данных предполагает соответствующее ограничение изучаемого частотного диапазона спектра. В частности, те колебания блеска квазара, которые выше трактовались как «вспышки», нельзя отождествлять с наиболее быстрыми вариациями потока - наблюдатели неоднократ-

но отмечали значительные флуктуации с характерным временем порядка нескольких десятков дней. Для изучения высокочастотной области спектра колебаний блеска 3C 273 необходимо привлечь все имеющиеся данные наблюдений и рассмотреть неравномерный временной ряд. Такое исследование родственного объекта - ядра сейфертовской галактики NGC 4151 - выполнено в работе В. Теребижа, А. Теребиж и Бирюкова (1989, см. также § 3.6).

§ 2.9. Взаимная спектральная плотность двух процессов

Введенные в § 1.5 понятия автоковариационной функции $\psi(\tau)$ и спектральной плотности $g(\nu)$ характеризуют внутреннюю структуру стационарного случайного процесса $\xi(t)$. Обобщением этих понятий, предназначенным для описания статистической связи между двумя различными процессами $\xi_1(t)$ и $\xi_2(t)$, являются *взаимная ковариационная функция* $\psi_{12}(\tau)$ и *взаимная спектральная плотность* $g_{12}(\nu)$. Примером задачи, требующей указанного обобщения, может служить нахождение свойств источника излучения путем сопоставления вариаций его светимости в двух различных диапазонах длин волн. Очевидно, установление зависимости между такими измерениями (в простейшем случае - просто запаздывания) дает важную информацию, как и изучение самого процесса колебаний светимости.

Взаимная корреляционная функция. Предположим, что $\xi_1(t)$ и $\xi_2(t)$ суть действительные стационарные в широком смысле случайные процессы с нулевыми средними и дисперсиями σ_1^2 и σ_2^2 соответственно. Обозначим автоковариационные функции процессов следующим образом:

$$\psi_{11}(\tau) = \langle \xi_1(t)\,\xi_1(t+\tau)\rangle, \quad \psi_{22}(\tau) = \langle \xi_2(t)\,\xi_2(t+\tau)\rangle. \qquad (2.9.1)$$

Как указывалось в § 1.5, эти функции действительны, симметричны и принимают при $\tau = 0$ значения, равные дисперсиям.

Легко видеть, что имеются две возможности ввести взаимную корреляционную функцию (ВКФ):

$$\psi_{12}(\tau) = \langle \xi_1(t)\,\xi_2(t+\tau)\rangle, \quad \psi_{21}(\tau) = \langle \xi_1(t+\tau)\,\xi_2(t)\rangle. \qquad (2.9.2)$$

однако ввиду равенства

$$\psi_{21}(\tau) = \psi_{12}(-\tau) \qquad (2.9.3)$$

достаточно изучать лишь одну из функций, скажем $\psi_{12}(\tau)$. В случае $\psi_{12}(\tau) \equiv 0$ процессы $\xi_1(t)$ и $\xi_2(t)$ являются некоррелированными; если рассматриваемые процессы относятся к гауссовским, то они при этом и независимы.

Один из наиболее простых примеров ВКФ следует из рассмотрения процесса $\xi_2(t)$, представляющего собой сдвинутый на интервал τ_0 процесс $\xi_1(t)$, т.е. $\xi_2(t) = \xi_1(t-\tau_0)$. Тогда

$$\psi_{12}(\tau) = \langle \xi_1(t)\xi_1(t+\tau-\tau_0)\rangle = \psi_{11}(\tau-\tau_0), \qquad (2.9.4)$$

т.е. ВКФ есть смещенная автоковариационная функция процесса $\xi_1(t)$.

Другой полезный пример доставляет рассмотрение линейной фильтрации типа (1.5.34). Именно, пусть $\xi_1(t)$ -процесс на входе линейного фильтра, с выхода которого поступает сигнал

$$\xi_2(t) = \int_{-\infty}^{\infty} h(t-t')\xi_1(t')dt' + n(t), \qquad (2.9.5)$$

где шум $n(t)$ не коррелирован с $\xi_1(t)$. Нетрудно убедиться, что в данном случае

$$\psi_{12}(\tau) = \int_{-\infty}^{\infty} h(\tau+\tau')\psi_{11}(\tau')d\tau'. \qquad (2.9.6)$$

Если импульсная характеристика фильтра $h(\tau) = \delta(\tau-\tau_0)$, то мы вновь приходим к формуле (4) для смещенного процесса.

В отличие от автоковариационной функции ВКФ не является, вообще говоря, четной. Как и всякую функцию, $\psi_{12}(\tau)$ можно разложить на сумму четной и нечетной составляющих:

$$\psi_{12}(\tau) = \psi_{12}^{+}(\tau) + \psi_{12}^{-}(\tau), \qquad (2.9.7)$$

где

$$\psi_{12}^{+}(\tau) = \frac{1}{2}\left[\psi_{12}(\tau) + \psi_{12}(-\tau)\right],$$
$$\psi_{12}^{-}(\tau) = \frac{1}{2}\left[\psi_{12}(\tau) - \psi_{12}(-\tau)\right]. \qquad (2.9.8)$$

Взаимная спектральная плотность. Согласно теореме Винера-Хинчина (1.5.13), автоковариационная функция $\psi(\tau)$ и спектральная плотность $g(\nu)$ стационарного случайного про-

цесса связаны парой преобразований Фурье. Исходя из взаимной корреляционной функции $\psi_{12}(\tau)$, можно аналогичным образом определить взаимную спектральную плотность:

$$g_{12}(\nu) = \int_{-\infty}^{\infty} e^{-i\cdot 2\pi\nu\tau}\psi_{12}(\tau)d\tau , \qquad (2.9.9')$$

так что

$$\psi_{12}(\tau) = \int_{-\infty}^{\infty} e^{i\cdot 2\pi\nu\tau} g_{12}(\nu)d\nu. \qquad (2.9.9'')$$

В отличие от $g(\nu)$, являющейся симметричной, положительно определенной функцией, взаимная плотность $g_{12}(\nu)$, вообще говоря, комплексна. Если воспользоваться в (9′) представлением (7) - (8), то мы получим

$$g_{12}(\nu) = g_{12}^{+}(\nu) - i\, g_{12}^{-}(\nu), \qquad (2.9.10)$$

где

$$g_{12}^{+}(\nu) = 2\int_{0}^{\infty} \cos(2\pi\nu\tau)\psi_{12}^{+}(\tau)d\tau,$$
$$g_{12}^{-}(\nu) = 2\int_{0}^{\infty} \sin(2\pi\nu\tau)\psi_{12}^{-}(\tau)d\tau \qquad (2.9.11)$$

- соответственно *коспектр* и *квадратурный спектр* двумерного случайного процесса $(\xi_1,\ \xi_2)$. Происхождение этих терминов связано с тем, что функция $g_{12}^{+}(\nu)$ учитывает корреляцию синфазных гармонических компонентов, а $g_{12}^{-}(\nu)$ - сдвинутых по фазе, или квадратурных, компонентов (см. Дженкинс и Ваттс, 1968, п.8.3.1).

Как и всякую комплексную величину, взаимную спектральную плотность можно представить не только в виде (10), но и следующим образом:

$$g_{12}(\nu) = \alpha_{12}(\nu)e^{i\varphi_{12}(\nu)}, \qquad (2.9.12)$$

где положительная функция $\alpha_{12}(\nu)$ называется *взаимным амплитудным спектром*, а функция $\varphi_{12}(\nu)$ $\propto\in[-\pi/2,\ \pi/2]$ - *взаимным фазовым спектром*. Очевидно, справедливы соотношения

$$\alpha_{12}(\nu) = \sqrt{\left[g_{12}^{+}(\nu)\right]^2 + \left[g_{12}^{-}(\nu)\right]^2}\,,$$

$$\varphi_{12}(\nu) = \operatorname{arctg}\left[-\frac{g_{12}^{-}(\nu)}{g_{12}^{+}(\nu)}\right]. \qquad (2.9.13)$$

Наконец, заметим, что $g_{12}^{+}(\nu)$ и $\alpha_{12}(\nu)$ являются четными функциями частоты, а $g_{12}^{-}(\nu)$ и $\varphi_{12}(\nu)$ - нечетными, вследствие чего $g_{12}^{-}(0) = 0$, $\varphi_{12}(0) = 0$.

Продолжим рассмотрение процесса линейной фильтрации (5) с точки зрения понятия взаимной спектральной плотности. Подставляя (6) в определение (9), находим

$$g_{12}(\nu) = H(\nu)\, g_{11}(\nu), \qquad (2.9.14)$$

где передаточная функция $H(\nu)$ есть, согласно (1.5.40), преобразование Фурье импульсной характеристики $h(\tau)$. Таким образом, мы пришли к фундаментальному для линейных систем выводу: взаимная спектральная плотность сигналов на выходе и входе линейного фильтра равна произведению спектральной плотности входного сигнала $g_{11}(\nu)$ и передаточной функции фильтра $H(\nu)$. Если последнюю записать в виде

$$H(\nu) = K(\nu)e^{i\psi(\nu)}, \qquad (2.9.15)$$

выделив коэффициент усиления фильтра $K(\nu) \geqslant 0$ и фазовый сдвиг фильтра $\psi(\nu)$, то из (12) и (14) следует

$$\alpha_{12}(\nu) = K(\nu)g_{11}(\nu), \quad \varphi_{12}(\nu) = \psi(\nu). \qquad (2.9.16)$$

Эти соотношения позволяют определить характеристики фильтра по измерению корреляции входного и выходного сигналов. В связи с (14) заметим, что собственная спектральная плотность выходного процесса (5) равна

$$g_{22}(\nu) = K^2(\nu)g_{11}(\nu) + g_{nn}(\nu), \qquad (2.9.17)$$

где $g_{nn}(\nu)$ - спектральная плотность шума. Формула (17) доказывается так же, как это было сделано при выводе

(1.5.41).

Простейший линейный фильтр - задержка на интервал τ_0 - характеризуется ВКФ (4) и передаточной функцией $H(\nu) = e^{-i\cdot 2\pi\nu\tau_0}$, так что для него

$$g_{22}(\nu) = g_{11}(\nu) + g_{nn}(\nu),\ g_{12}(\nu) = e^{-i\cdot 2\pi\nu\tau_0} g_{11}(\nu),$$
$$\alpha_{12}(\nu) = g_{11}(\nu),\quad \varphi_{12}(\nu) = -2\pi\nu\tau_0 \qquad (2.9.18)$$

- взаимная спектральная плотность включает лишь фазовый сдвиг, пропорциональный частоте.

Когерентность. Введение взаимной спектральной плотности $g_{12}(\nu)$ имело целью характеризовать зависимость между отдельными спектральными компонентами двух случайных процессов $\xi_1(t)$ и $\xi_2(t)$. В частности, взаимный амплитудный спектр $\alpha_{12}(\nu)$ в (12) служит мерой связи спектральных плотностей процессов, а фазовый спектр $\varphi_{12}(\nu)$ - мерой временного запаздывания. Вместе с тем следует отметить, что непосредственное использование $\alpha_{12}(\nu)$ сопряжено с неудобством, аналогичным возникающему при использовании коэффициента автоковариации для обычных случайных величин ξ_1 и ξ_2: этот коэффициент зависит от диапазона значений, принимаемых случайными величинами. Для последних мерой линейной связи, свободной от указанного недостатка, является квадрат коэффициента корреляции

$$\rho^2 = \frac{\langle \xi_1 \xi_2 \rangle^2}{\sigma_1^2\,\sigma_2^2}, \qquad (2.9.19)$$

значения которого заключены в пределах от 0 до 1. При изучении связи между процессами $\xi_1(t)$ и $\xi_2(t)$ на некоторой частоте ν аналогом (19) является соотношение

$$C(\nu) = \frac{|g_{12}(\nu)|^2}{g_{11}(\nu)g_{22}(\nu)} = \frac{\alpha_{12}^2(\nu)}{g_{11}(\nu)g_{22}(\nu)}, \qquad (2.9.20)$$

где *когерентность* $C(\nu)$ также принимает значения в промежутке [0, 1]. Сходство двух последних соотношений станет особенно ясным, если учесть теорему Рэлея (1.5.14) для случайных функций и формулы (2) и (9″) настоящего параграфа.

Найдем в качестве примера величину $C(\nu)$ для случая линейной фильтрации процесса вида (5). Учитывая в (20) формулы (16) и (17), получаем выражение

$$C(\nu) = \frac{K^2(\nu)g_{11}(\nu)}{K^2(\nu)g_{11}(\nu) + g_{nn}(\nu)}, \qquad (2.9.21)$$

включающее только характеристики входного процесса и фильтра. Знаменатель здесь представляет собой спектральную плотность $g_{22}(\nu)$ выходного процесса, а числитель - ту ее часть, которая обусловлена входным процессом, т.е. $C(\nu)$ есть отношение сигнала S на выходе фильтра к сумме сигнала и шума $S + N$. Естественно, что при $S \ll N$ когерентность близка к 0, а при $S \gg N$ - к 1. Таким образом, когерентность для линейного фильтра является мерой относительной мощности сигнала на данной частоте. К этому же выводу приводит рассмотрение соотношения

$$g_{22}(\nu) = \frac{g_{nn}(\nu)}{1 - C(\nu)}, \qquad (2.9.22)$$

следующего из (2:) и (17).

Дискретность отсчетов. В общем случае наблюдатель располагает N_1 отсчетами некоторого процесса $y_1(t)$, полученными в моменты времени $(t_0, t_1, \ldots, t_{N_1-1})$, и N_2 отсчетами другого процесса $y_2(t)$ для моментов $(t'_0, t'_1, \ldots, t'_{N_2-1})$. Предположим сначала, что оба ряда равномерны с шагом по времени Δt и содержат одинаковое число отсчетов $N_1 = N_2 = N$. По аналогии с (2.3.7) в качестве оценки взаимной спектральной плотности $g_{12}(\nu)$ рассмотрим величину $N\Delta t P_{12}(\nu)$, где *взаимная периодограмма* по определению равна

$$P_{12}(\nu) = \frac{1}{N}\sum_{k=0}^{N-1} e^{-i\cdot 2\pi\nu\Delta t k} y^{(0)}_{1,k} \cdot \frac{1}{N}\sum_{l=0}^{N-1} e^{i\cdot 2\pi\nu\Delta t l} y^{(0)}_{2,l}. \qquad (2.9.23)$$

Производя здесь фактическое перемножение рядов, находим

$$P_{12}(\nu) = \frac{1}{N}\left[c_0 + \sum_{n=1}^{N-1}\cos(2\pi\nu\Delta t n)(c'_n + c''_n) + i\sum_{n=1}^{N-1}\sin(2\pi\nu\Delta t n)(c'_n - c''_n)\right], \qquad (2.9.24)$$

где выборочные коэффициенты взаимной ковариации

$$c'_n = \frac{1}{N}\sum_{k=0}^{N-1-n} y^{(0)}_{1,k}\, y^{(0)}_{2,k+n},$$

$$c''_n = \frac{1}{N}\sum_{k=0}^{N-1-n} y^{(0)}_{1,k+n}\, y^{(0)}_{2,k}, \quad n = 0, 1, \ldots, N-1, \tag{2.9.25}$$

и $c_0 = c'_0 = c''_0$. Если отсчеты выбраны из одного и того же ряда, т.е., $y_1 \equiv y_2$, то $c'_n = c''_n$, и мы приходим к известному выражению (2.1.26) периодограммы через выборочную автоковариационную функцию. Для различных рядов $P_{12}(\nu)$ содержит мнимое слагаемое, отвечающее квадратурному спектру (10), (11).

Статистические свойства взаимной $P_{12}(\nu)$ и обычной $P_N(\nu)$ периодограммы весьма сходны. В частности, величина $N\Delta t P_{12}(\nu)$ не является состоятельной оценкой истинной спектральной плотности $g_{12}(\nu)$, т.е. дисперсия $N\Delta t P_{12}(\nu)$ не стремится к нулю при неограниченном увеличении N. Поэтому нахождение состоятельных оценок $g_{12}(\nu)$ должно опираться на те же принципы, которые рассматривались выше в связи с оценкой $g(\nu)$. Подробное изложение относящихся сюда вопросов можно найти в книге Дженкинса и Ваттса (1968).

Расширение понятия взаимной периодограммы на упомянутый в начале данного пункта общий случай можно произвести следующим образом:

$$P_{12}(\nu) = \frac{1}{N_1 N_2} F_{N_1}(\nu) F^*_{N_2}(\nu), \tag{2.9.26}$$

где

$$F_{N_1}(\nu) = \sum_{k=0}^{N_1-1} e^{-i\cdot 2\pi\nu t_k}\, y^{(0)}_1(t_k),$$

$$F_{N_2}(\nu) = \sum_{l=0}^{N_2-1} e^{-i\cdot 2\pi\nu t'_l}\, y^{(0)}_2(t'_l). \tag{2.9.27}$$

Для выяснения особенностей этого определения желательно применить его к различным модельным задачам.

ГЛАВА 3

СПЕКТРАЛЬНАЯ ТЕОРИЯ НЕРАВНОМЕРНЫХ ВРЕМЕННЫХ РЯДОВ

Строго равномерное распределение моментов наблюдения временного процесса является идеализацией, не достижимой на практике даже в том случае, когда поставлена специальная задача получить равномерный ряд. Во-первых, всегда существуют ошибки в регистрации моментов времени или пространственной координаты при формировании совокупности отсчетов процесса. Кроме того, в астрономии существенная неравномерность временных рядов обусловлена большой длительностью изучаемых процессов, непостоянством погодных условий, загруженностью телескопов и другими факторами, не зависящими от исследователя.

Важность практических приложений лишь отчасти поясняет причину особого интереса, который вызывает изучение неравномерных временных рядов. Дело в том, что должна быть выяснена проблема устойчивости результатов, полученных для строгой последовательности $t_k^{(e)} = k\Delta t$, $k = 0, 1, \ldots, N-1$, при малых отклонениях от равномерности ряда. Различные результаты могут быть в разной степени устойчивы в указанном смысле, и лишь после соответствующего исследования можно полагаться на конкретные выводы теории равномерных временных рядов.

§ 3.1. LS_m-спектр

Как было показано в гл. 2, основой для построения оценок истинной спектральной плотности процесса $g(\nu)$ в случае равномерных временных рядов может служить периодограмма

Шустера $P_N(\nu)$, определенная соотношениями (1.1.11) и (1.1.13). Вполне естественно, что при необходимости анализировать неравномерные ряды исследователи издавна пользовались простым обобщением классической периодограммы, сводящимся к замене последовательности $t_k^{(e)} = k\Delta t$ произвольно распределенной совокупностью моментов наблюдения $\{t_k\}$, как это и было сделано в § 1.6 (см. (1.6.12)). Целесообразность такого обобщения при анализе неравномерных рядов подтверждают многочисленные практические приложения, однако указанный подход не имеет ни обоснования, ни развитой теории статистических оценок. Между тем очевидно, что для произвольной совокупности $\{t_k\}$ значения периодограммы $P_N(\nu)$ не подчиняются экспоненциальному распределению (2.3.25).

Итак, первоочередной задачей при изучении неравномерных рядов является отыскание такого обобщения периодограммы Шустера, которое имело бы достаточно глубокое обоснование и сохранило привычные свойства последней. Искомое обобщение - так называемый LS-*спектр* $\mathcal{P}_N(\nu)$ (от Least Squares) - было предложено Барнингом (1963), Ванишеком (1971) и Ломбом (1976). Весьма существенно, что для многих реальных ситуаций отличие $\mathcal{P}_N(\nu)$ от периодограммы $P_N(\nu)$, формально определенной для неравномерной последовательности $\{t_K\}$, оказалось незначительным. Это обстоятельство поясняет эффективность применений $P_N(\nu)$ в ранних исследованиях. Вместе с тем необходимо подчеркнуть, что только использование LS-спектра обеспечивает уверенность при анализе спектра неравномерного ряда и оценке значимости максимумов спектральной плотности.

Приводимый ниже вывод LS-спектра несколько отличается от оригинального, поскольку мы считали полезным уделить больше внимания интуитивному обоснованию применяемых процедур.

Модель. Рассмотрим модель, согласно которой наблюдаемые значения процесса $\{y_k\}$ представляют собой сумму некоторой периодической функции $f(t_k)$ с периодом P_0 и случайных величин (шума) $\xi_k \equiv \xi(t_k)$:

$$y_k = f(t_k) + \xi_k, \quad k = 0, 1, \ldots, N-1, \qquad (3.1.1)$$

причем моменты наблюдения $\{t_k\}$ произвольно распределены по оси времени. Запишем разложение Фурье функции $f(t)$ в виде

$$f(t) = \sum_{n=1}^{\infty} \left[\bar{a}_n \cos(2\pi n \nu_0 t) + \bar{b}_n \sin(2\pi n \nu_0 t) \right], \qquad (3.1.2)$$

где $\nu_0 = P_0^{-1}$ и мы приняли для простоты среднее значение $f(t)$ по периоду равным нулю. Предположим далее, что ξ_k суть попарно независимые случайные величины с нулевым средним, дисперсией σ^2 и нормальной плотностью распределения

$$p(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-x^2/(2\sigma^2)}, \quad -\infty < x < \infty. \qquad (3.1.3)$$

Перед нами стоит задача найти по наблюдаемой совокупности $\{t_k, y_k\}$ некоторые «хорошие» оценки неизвестных параметров $(\bar{a}_n, \bar{b}_n, \nu_0)$ и тем самым выделить из наблюдаемого ряда $\{y_k\}$ периодический компонент $\{f(t_k)\}$. Термин «хорошие» означает, что искомые оценки $(\hat{a}_n, \hat{b}_n, \hat{\nu}_0)$ должны быть *несмещенными* (среднее значение оценки равно истинному значению) и *состоятельными* (при $N \to \infty$ оценки все более должны группироваться около истинных значений). Для решения задачи применим *метод максимального правдоподобия*, приводящий именно к таким оценкам (см., например, Крамер 1975).

Оценки максимального правдоподобия. Обозначим через L плотность вероятности получить ту совокупность значений процесса, которая наблюдалась в действительности. Поскольку величины $\{\xi_k\}$ по предположению независимы и распределены нормально, из (1) и (3) следует

$$L = \left(\frac{1}{\sigma \sqrt{2\pi}} \right)^N \exp \left\{ -\frac{1}{2\sigma^2} \sum_{k=0}^{N-1} \left[y_k - f(t_k) \right]^2 \right\}. \qquad (3.1.4)$$

Согласно методу максимального правдоподобия, в качестве оценок неизвестных параметров следует принять такие значения $(\hat{a}_n, \hat{b}_n, \hat{\nu}_0)$, при которых функция правдоподобия L достигает максимума. Из (4) видно, что максимум L достигается в том случае, когда сумма в показателе экспоненты минимальна. Мы приходим к известному результату: оценки максимального правдоподобия совпадают для нормального случайного процесса с оценками согласно

методу наименьших квадратов, и для их нахождения следует использовать условие

$$\sum_{k=0}^{N-1}\left[y_k - f(t_k)\right]^2 = \min. \qquad (3.1.5)$$

Фактическое применение (5) производится следующим образом. Введем среднеквадратическое уклонение

$$\Delta_N^2(\nu) = \frac{1}{N}\sum_{k=0}^{N-1}\left\{y_k - \sum_{n=1}^{\infty}\left[a_n\cos(2\pi n\nu t) + b_n\sin(2\pi n\nu t)\right]\right\}^2 \qquad (3.1.6)$$

с произвольными параметрами (a_n, b_n, ν) и будем искать минимум $\Delta_N^2(\nu)$, исходя из условий

$$\frac{\partial\Delta_N^2}{\partial\nu} = 0, \qquad \frac{\partial\Delta_n^2}{\partial a_n} = 0, \qquad \frac{\partial\Delta_n^2}{\partial b_n} = 0. \qquad (3.1.7)$$

Решение уравнений (7) - так называемой *системы нормальных уравнений* - и дает искомые оценки ($\hat{a}_n$, $\hat{b}_n$, $\hat{\nu}_0$). На практике приходится ограничиваться конечным числом m слагаемых во внутренней сумме (6), так что (7) представляет собой систему $2m + 1$ уравнений, но даже при небольших значениях m система весьма громоздка. Поэтому разыскание глобального минимума величины $\Delta_N^2(\nu)$ можно проводить не путем решения полной системы нормальных уравнений, а постепенно, путем последовательного подбора гармоник, доставляющих (6) локальные минимумы. Поскольку в формировании наблюдаемых значений $\{y_k\}$ внесли вклад все гармоники (2) функции $f(t)$, из (6) ясно, что каждый раз, когда при некотором n параметры a_n, b_n и частота ν становятся близкими к искомым оценкам $\hat{a}_n$, $\hat{b}_n$, $\hat{\nu}_0$, значение $\Delta_N^2(\nu)$ уменьшается по сравнению с ее значением при произвольных коэффициентах. В этом последнем случае Δ_N^2 приблизительно равно тому его значению $\Delta_{N,0}^2$, которое соответствует тождественному нулю в качестве аппроксимирующей функции:

$$\Delta_{N,0}^2 = \frac{1}{N}\sum_{k=0}^{N-1} y_k^2 , \qquad (3.1.8)$$

т.е. просто выборочной дисперсии процесса. При этом уменьшение дисперсии $\Delta_{N,0}^2 - \Delta_N^2(\nu)$ характеризует интенсивность

рассматриваемого гармонического колебания в процессе $\{y_k\}$.

Формула Барнинга. Приведенный анализ подсказывает дальнейший ход рассуждений при введении понятия LS-спектра. Вместо того чтобы использовать в (6) при минимизации $\Delta_N^2(\nu)$ полигармонический процесс, оставим лишь одно гармоническое колебание

$$\begin{aligned} \tilde{f}_\nu(t) &= A(\nu)\cos(2\pi\nu t - \varphi) = \\ &= a(\nu)\cos(2\pi\nu t) + b(\nu)\sin(2\pi\nu t), \\ a(\nu) &= A(\nu)\cos\varphi, \quad b(\nu) = A(\nu)\sin\varphi, \end{aligned} \tag{3.1.9}$$

и при каждом значении частоты ν подберем функции $a(\nu)$ и $b(\nu)$ так, чтобы значение

$$\tilde{\Delta}_N^2(\nu) = \frac{1}{N}\sum_{k=0}^{N-1}\left[y_k - \tilde{f}_\nu(t_k)\right] \tag{3.1.10}$$

было минимально. Как отмечалось, уменьшение дисперсии является мерой интенсивности колебания $\tilde{f}_\nu(t)$ в процессе $y(t)$. По смыслу это уменьшение соответствует общепринятому толкованию спектра мощности, так что мы определим LS-спектр $\mathcal{P}_N(\nu)$ соотношением

$$\mathcal{P}_N(\nu) \equiv \frac{1}{2}\left\{\Delta_{N,0}^2 - \left[\tilde{\Delta}_N^2(\nu)\right]_{\min}\right\}, \tag{3.1.11}$$

где множитель 1/2 введен для удобства сопоставления с классической периодограммой $P_N(\nu)$.

Найдем явное выражение для $\mathcal{P}_N(\nu)$. Из (9) и (10) следует

$$\tilde{\Delta}_N^2(\nu) = \frac{1}{N}\sum_{k=0}^{N-1}\left[y_k - a(\nu)\cos(2\pi\nu t_k) - b(\nu)\sin(2\pi\nu t_k)\right]^2. \tag{3.1.12}$$

Приравнивая производные от (12) по a и b к нулю, а также вводя обозначения

$$\begin{aligned} C_N(\nu) &= \frac{2}{N}\sum_{k=0}^{N-1}\cos^2(2\pi\nu t_k) = 1 + \frac{1}{N}\sum_{k=0}^{N-1}\cos(4\pi\nu t_k), \\ S_N(\nu) &= \frac{1}{N}\sum_{k=0}^{N-1}\sin(4\pi\nu t_k), \\ U_N(\nu) &= \frac{1}{N}\sum_{k=0}^{N-1}\cos(2\pi\nu t_k)y_k, \end{aligned} \tag{3.1.13}$$

$$V_N(\nu) = \frac{1}{N}\sum_{k=0}^{N-1} \sin(2\pi\nu t_k) y_k,$$

получаем нормальную систему уравнений

$$\begin{aligned} aC_N(\nu) + bS_N(\nu) &= 2U_N(\nu), \\ aS_N(\nu) + b[2 - C_N(\nu)] &= 2V_N(\nu). \end{aligned} \qquad (3.1.14)$$

Решение (14), минимизирующее $\tilde{\Delta}^2_N(\nu)$, имеет вид:

$$\begin{aligned} \hat{a}(\nu) &= 2\,\frac{(2-C_N)U_N - S_N V_N}{C_N(2-C_N)-S_N^2}, \\ \hat{b}(\nu) &= 2\,\frac{C_N V_N - S_N U_N}{C_N(2-C_N)- S_N^2}. \end{aligned} \qquad (3.1.15)$$

Само минимальное значение $[\tilde{\Delta}^2_N(\nu)]_{\min}$ находится путем подстановки (15) в выражение (12); нетрудно показать, что

$$\left[\tilde{\Delta}^2_N(\nu)\right]_{\min} = \Delta^2_{N,0} - \hat{a}(\nu)U_N(\nu) - \hat{b}(\nu)V_N(\nu). \qquad (3.1.16)$$

Наконец, из (16) и (11) следует явное выражение для LS-спектра, найденное Ф. Барнингом (1963):

$$\mathcal{P}_N(\nu) = \frac{(2-C_N)U_N^2 + C_N V_N^2 - 2S_N U_N V_N}{C_N(2 - C_N) - S_N^2}. \qquad (3.1.17)$$

где функции $C_N(\nu)$, $S_N(\nu)$, $U_N(\nu)$ и $V_N(\nu)$ определены соотношениями (13). Последние две из упомянутых функций совпадают с обычными дискретными преобразованиями Фурье типа (1.1.11), но теперь $\{t_k\}$ - не обязательно равноотстоящие друг от друга моменты времени.

Представление Ломба. В работе Ломба (1976) найдена другая форма LS-спектра, иногда более удобная при теоретическом рассмотрении или расчетах. Ломб обратил внимание, что смешанный коэффициент S_N в системе (14) можно уничтожить, если

ввести запаздывание

$$\tau(\nu) = \frac{1}{4\pi\nu}\,\mathrm{arctg}\,\frac{\sum_{k=0}^{N-1}\sin(4\pi\nu t_k)}{\sum_{k=0}^{N-1}\cos(4\pi\nu t_k)} \qquad (3.1.18)$$

и сдвинуть соответствующим образом начальную точку отсчета временной оси. Тогда вместо (13) мы имеем

$$\begin{aligned} C_N(\nu,\tau) &= 1 + \frac{1}{N}\sum_{k=0}^{N-1}\cos\,[4\pi\nu(t_k - \tau)], \\ S_N(\nu,\tau) &= \frac{1}{N}\sum_{k=0}^{N-1}\sin\,[4\pi\nu(t_k - \tau)] = 0, \\ U_N(\nu,\tau) &= \frac{1}{N}\sum_{k=0}^{N-1}\cos\,[2\pi\nu(t_k - \tau)]y_k, \\ V_N(\nu,\tau) &= \frac{1}{N}\sum_{k=0}^{N-1}\sin\,[2\pi\nu(t_k - \tau)]y_k, \end{aligned} \qquad (3.1.19)$$

а выражение (17) принимает вид

$$\mathcal{P}_N(\nu) = \frac{U_N^2(\nu,\tau)}{C_N(\nu,\tau)} + \frac{V_N^2(\nu,\tau)}{2-C_N(\nu,\tau)}\,. \qquad (3.1.20)$$

Это и есть представление Ломба для LS-спектра, весьма сходное с определением (1.6.12) классической периодограммы $P_N(\nu)$.

Если число отсчетов временного ряда N задано окончательно, то вычисления, вероятно, целесообразно проводить по формулам Ломба (18) - (20). Однако при последовательном анализе все более длинного ряда придется каждый раз вновь вычислять $\tau(\nu)$ и суммы (19), так что в этом случае лучше вернуться к оригинальному выражению (17) для LS-спектра.

Равномерный ряд. Периодограмма Шустера представляет собой квадрат коэффициента корреляции между временным рядом и гармоническим колебанием (с точностью до постоянного множителя), LS-спектр характеризует качество аппроксимации ряда гармоническим колебанием в смысле среднеквадратического от-

клонения. Мы покажем сейчас, что эти, казалось бы, разные подходы оказываются эквивалентными для не слишком коротких равномерных рядов.

Итак, пусть $t_k = k\Delta t$, $k = 0, 1, \ldots, N-1$. Пользуясь формулами (П. 1.1) и (П. 1.2), получаем:

$$\begin{cases} C_N(\nu) = 1 + \dfrac{1}{2N}\left\{1 + \dfrac{\sin[(2N-1)2\pi\nu\Delta t]}{\sin(2\pi\nu\Delta t)}\right\}, \\ S_N(\nu) = \dfrac{1}{N}\dfrac{\sin(N2\pi\nu\Delta t)\sin[(N-1)2\pi\nu\Delta t]}{\sin(2\pi\nu\Delta t)}. \end{cases} \qquad (3.1.21)$$

Во внутренних точках интервала Найквиста $0 \leqslant \nu \leqslant (2\Delta t)^{-1}$ коэффициент $C_N(\nu)$ мало отличается от единицы, а $S_N(\nu)$ - от нуля:

$$|1 - C_N(\nu)| \sim N^{-1}, \quad |S_N(\nu)| \sim N^{-1}; \qquad (3.1.22)$$

на границах указанного интервала $C_N = 2$. При $N \gg 1$ можно считать $C_N(\nu)=1$, $S_N(\nu)=0$, и выражение (17) значительно упрощается:

$$\mathcal{P}_N(\nu) \simeq U_N^2(\nu) + V_N^2(\nu) \equiv P_N(\nu), \quad N \gg 1. \qquad (3.1.23)$$

Это соотношение доказывает высказанное выше утверждение. Таким образом, LS-спектр может рассматриваться как естественное расширение Фурье-анализа на случай неравномерных временных рядов; при определенных условиях, сформулированных в начале данного параграфа, LS-спектр оказывается и оптимальным таким обобщением периодограммы.

Интерпретация классической периодограммы Шустера равномерного ряда как меры уменьшения дисперсии ряда при наилучшей аппроксимации его синусоидальным колебанием весьма полезна для практики. Согласно (11), $P_N(\nu)$ составляет половину уменьшения дисперсии на частоте ν.

Следует заметить, что найденная при численных экспериментах и анализе реальных неравномерных рядов близость обобщенной периодограммы Шустера и LS-спектра не исследована еще должным образом. В частности, интересно изучить распределения моментов наблюдения, для которых не выполняются условия (22).

Статистика LS-спектра. Для оценки значимости максимумов LS-спектра необходимо знать, как часто такие максимумы обнаруживаются при анализе чисто случайного процесса. В § 2.3 было показано, что для равномерного временного ряда, составленного из отсчетов белого шума, значения $P_N(\nu)$ на некоторой частоте характеризуются экспоненциальным распределением. Можно доказать, что экспоненциальный закон не сохраняется при формальном обобщении периодограммы (1.6.12) на случай произвольного распределения моментов наблюдения. Строго говоря, это обстоятельство в значительной степени обесценивает подход, основанный на вычислении $P_N(\nu)$ согласно (1.6.12), однако на практике следует учитывать два важных результата: 1) Ломб (1976) доказал, что значения LS-спектра для произвольной совокупности $\{t_k\}$ распределены экспоненциально; 2) по-видимому, существенного отличия между $P_N(\nu)$ и $\mathcal{P}_N(\nu)$ можно ожидать лишь при «патологических» распределениях моментов наблюдения (см. Ломб, 1976; Скаргл, 1982; Хорн и Балиунас, 1986).

Таким образом, обычно с известной долей уверенности можно полагать, что распределение значений $P_N(\nu)$ неравномерного ряда близко к экспоненциальному и, следовательно, можно применять статистические критерии, изложенные в § 2.3. Однако надежный (и лишь немногим более трудоемкий) путь состоит в нахождении LS-спектра $\mathcal{P}_N(\nu)$ изучаемого процесса по формулам Барнинга (17) или Ломба (20). При этом единственной проблемой в применении стандартной методики остается оценка количества независимых частот N_i - показателя степени в распределении Уокера

$$Q^{(W)}_{N_i}(x) = 1 - (1 - e^{-x})^{N_i}, \quad 0 \leqslant x < \infty, \qquad (3.1.24)$$

являющемся обобщением (2.3.31) на случай неравномерного ряда. Численные эксперименты Хорна и Балиунас (1986) с некоторыми видами неравномерности показали, что значения N_i оказываются порядка длины ряда N. Вообще же значимость максимумов $\mathcal{P}_N(\nu)$ для конкретного ряда можно оценить путем

многократного численного моделирования белого шума с теми же средним значением, дисперсией и распределением моментов наблюдения $\{t_k\}$, что и для изучаемого ряда, а еще лучше - путем многократного случайного перемешивания отсчетов при сохранении $\{t_k\}$ с вычислением для каждого искусственного временного ряда LS-спектра или периодограммы.

LS_m- спектр. Расчет $\mathcal{P}_N(\nu)$ по своей сущности аналогичен исследованию спектра звуковых колебаний при помощи набора пробных камертонов, каждый из которых реагирует на чистый тон - синусоидальное колебание (или при помощи одного резонатора с перестраиваемой частотой). Однако может случиться так, что в спектре процесса нет резко выделяющейся линии, например, искомое колебание имеет сложную форму, для которой гармоники Фурье мало отличаются по амплитуде. Если вдобавок ряд сильно зашумлен, трудно надеяться, что LS-спектр выделит периодический компонент. Ясно, что в этих условиях более эффективным будет вычисление функции $\Delta_N^2(\nu)$ из (6) для интересующего нас диапазона частот, т.е. одновременное исследование спектра при помощи набора резонаторов с гармонической последовательностью частот.

Ограничиваясь конечным числом гармоник и вводя для общности смещение A_0, перепишем (6) в виде

$$\Delta^2_{N,m}(\nu) = \frac{1}{N}\sum_{k=0}^{N-1}\Big\{Y_k - A_0 - \sum_{n=1}^{m}\Big[a_n\cos(2\pi n\nu t) + b_n\sin(2\pi n\nu t)\Big]\Big\}^2. \qquad (3.1.25)$$

Для каждого значения частоты функцию $\Delta^2_{N,m}(\nu)$ следует минимизировать, т.е. найти такие значения коэффициентов $A_0(\nu)$ и $\{a_n(\nu),\ b_n(\nu)\}$, при которых остаточная дисперсия (25) достигает минимального значения $[\Delta^2_{N,m}(\nu)]_{\min}$. Затем образуем по типу (11) разность

$$\mathcal{P}_{N,m}(\nu) \equiv \frac{1}{2}\left\{\frac{1}{N}\sum_{k=0}^{N-1} y_k^2 - \left[\Delta^2_{N,m}(\nu)\right]_{\min}\right\}, \qquad (3.1.26)$$

которая и представляет собой обобщение LS-спектра на случай полигармонического приближения. Мы назовем функцию $\mathcal{P}_{N,m}(\nu)$ для краткости LS_m-*спектром*, сохраняя обозначение LS ($\equiv LS_1$)-спектр для функции $\mathcal{P}_N(\nu)$.

Подход, близкий к описанному выше, давно используется в астрономии. Так, Пельт (1980) детально рассматривает методику вычисления функции $[\Delta^2_{N,m}(\nu)]_{\min}$, названной им D_2-*статистикой*. Следует, впрочем, заметить, что к понятию собственно спектра мощности привело бы дальнейшее применение соотношения (26).

Нахождение оптимальных значений коэффициентов в (25) требует решения системы уравнений, практически совпадающей с (3.3.9). При этом возникает вопрос о наилучшем выборе порядка m, обсуждение которого мы отложим до § 3.3. Вообще нужно сказать, что вычисление LS_m-спектра является трудоемкой задачей, которую часто приходится решать лишь для сравнительно небольшого интервала частот после изучения ряда простыми и быстрыми способами. Тем более трудоемким для численных расчетов представляется теоретически очевидное обобщение спектра, когда он рассматривается как отклик на несколько разных периодических колебаний.

§ 3.2. Случайная последовательность моментов наблюдения

До сих пор единственным случайным фактором, входившим в наше рассмотрение, был стохастический компонент $\{\xi(t_k)\}$ наблюдаемого временного ряда $\{y(t_k)\}$. Теперь мы предположим, что и совокупность моментов наблюдения $\{t_k\}$ подчиняется некоторому вероятностному распределению. В простейшем случае вероятностный характер $\{t_k\}$ может быть обусловлен просто ошибками измерения времени (или координаты), так что

$$t_k = t_k^{(e)} + \varepsilon_k \Delta t\,, \quad k = 0,\ 1,\ \ldots,\ N-1, \qquad (3.2.1)$$

где $t_k^{(e)} = k\Delta t$ - строго равномерная сетка моментов регистрации процесса, а ε_k - случайные ошибки измерения, выраженные в единицах Δt. В других случаях эффектами, связанными с

ошибками $\{\varepsilon_k\}$, можно пренебречь, а случайный характер $\{t_k\}$ ввести как приближение, описывающее типичную совокупность моментов наблюдения. Такой подход позволяет выяснить некоторые общие свойства неравномерных рядов, не зависящие от конкретной реализации $\{t_k\}$.

Ниже выведено соотношение (7), связывающее спектральную плотность процесса $g(\nu)$ со средней периодограммойпри достаточно широких предположениях, а затем даны приложения общих результатов к двум в определенном смысле противоположным случаям: а) моменты наблюдения случайно распределены на большом промежутке $[0,T]$, б) эти моменты случайно отклоняются от заданной наперед равномерной сетки $t_k^{(e)}$. В обоих случаях средняя периодограмма весьма просто выражается через характеристическую функцию случайных моментов наблюдения, а выводы допускают наглядное истолкование. Возможность использования периодограммы вместо LS-спектра основана на результатах, описанных в предыдущем параграфе.

Периодограмма Шустера случайного временного ряда. Пусть $\xi(t)$ - центрированный стационарный случайный процесс со спектральной плотностью $g(\nu)$, $\xi_k = \xi(t_k)$, $k = 0, 1, \ldots, N-1$, - совокупность его зарегистрированных отсчетов. Как и в (1.6.28), определим периодограмму временного ряда соотношением

$$P_N(\nu) = \left| \frac{1}{N} \sum_{k=0}^{N-1} e^{-i\cdot 2\pi\nu t_k} \xi(t_k) \right|^2 = \tag{3.2.2}$$

$$= \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} e^{-i\cdot 2\pi\nu(t_k - t_l)} \xi(t_k)\xi(t_l),$$

однако здесь случайный характер $P_N(\nu)$ связан не только со случайностью значений $\{\xi(t_k)\}$, но и со случайным выбором самих моментов наблюдения. Считая пока эти последние фиксированными, произведем усреднение (2) по ординатам $\xi(t_k)$. Принимая во внимание теорему Винера-Хинчина

$$\langle \xi(t_k)\xi(t_l) \rangle = \psi(t_k - t_l) = \int_{-\infty}^{\infty} e^{i\cdot 2\pi\nu'(t_k - t_l)} g(\nu')\,d\nu' \tag{3.2.3}$$

(см. § 1.5), получаем

$$\langle P_N(\nu)\rangle = \int_{-\infty}^{\infty} g(\nu')\left[\frac{1}{N^2}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} e^{-i\cdot 2\pi(\nu-\nu')(t_k-t_l)}\right]d\nu'. \quad (3.2.4)$$

усредним теперь (4) по совокупности моментов $\{t_k\}$; полагая

$$R_N(\nu) \equiv \left\langle \frac{1}{N^2}\sum_{l=0}^{N-1}\sum_{l=0}^{N-1} e^{-i\cdot 2\pi\nu(t_k-t_l)}\right\rangle = \quad (3.2.5)$$

$$= \frac{1}{N} + \frac{1}{N^2}\sum_{\substack{k,l=0\\ k\neq l}}^{N-1} \langle e^{-i\cdot 2\pi\nu(t_k-t_l)}\rangle,$$

$$R_N(0) = 1, \quad (3.2.6)$$

находим искомое соотношение между средней периодограммой и спектральной плотностью процесса:

$$\langle\langle P_N(\nu)\rangle\rangle = \int_{-\infty}^{\infty} R_N(\nu-\nu')g(\nu')d\nu'. \quad (3.2.7)$$

Очевидно, при детерминированной совокупности $t_k = t_k^{(0)}$, когда плотность совместного распределения $\{t_k\}$ является произведением δ-функций, усреднение в (5) выполняется весьма просто и

$$R_N(\nu) = |W_N^{(0)}(\nu)|^2 = \left|\frac{1}{N}\sum_{k=0}^{N-1} e^{-i\cdot 2\pi\nu t_k^{(0)}}\right|^2 \quad (3.2.8)$$

- мы вновь приходим к формуле (1.6.31). Если временной ряд вдобавок равномерен, т.е. $t_k^{(0)} = t_k^{(e)} = k\Delta t$, то

$$R_N(\nu) = |\phi_N(\nu\,\Delta t)|^2, \quad (3.2.9)$$

где функция $\phi_N(x)$ определена выражением (2.1.4).

Рассмотрим некоторые конкретные приложения общих результатов, полученных в данном пункте.

Независимые, одинаково распределенные t_k. По-видимому, наиболее простой пример доставляет нам рассмотрение случая, когда все моменты наблюдения суть одинаково распределенные на промежутке $[0,T]$ случайные величины, попарно независимые между собой. Обозначим через $p(t)$ плотность распределения вероятностей величины t_k, так что

$$\Pr[t < t_k \leqslant t+dt] = p(t)dt, \quad 0 \leqslant t \leqslant T, \quad (3.2.10)$$

и через $H(\nu)$ - *характеристическую функцию* t_k, являющуюся по определению преобразованием Фурье плотности распределения:

$$H(\nu) = \int_0^T e^{-i\cdot 2\pi\nu t}\, p(t)dt = \langle e^{-i\cdot 2\pi\nu t_k}\rangle, \qquad (3.2.11)$$

$$H(0) = 1 . \qquad (3.2.12)$$

Тогда из (5) находим:

$$R_N(\nu) = \frac{1}{N^2}\left[N + \sum_{\substack{k,l=0\\ k\neq l}}^{N-1} \langle e^{-i\cdot 2\pi\nu t_k}\rangle\langle e^{i\cdot 2\pi\nu t_l}\rangle\right] =$$

$$= \frac{1}{N^2}\left[N + N(N-1)H(\nu)H^*(\nu)\right]. \qquad (3.2.13)$$

Окончательно,

$$R_N(\nu) = \frac{1}{N} + \left(1-\frac{1}{N}\right)|H(\nu)|^2. \qquad (3.2.14)$$

Таким образом, в данном случае ядро основного соотношения (7) весьма просто выражается через характеристическую функцию случайного момента наблюдения $H(\nu)$. При $N \gg 1$ ядро вообще не зависит от числа измерений,

$$R_N(\nu) \simeq |H(\nu)|^2, \qquad N \gg 1, \qquad (3.2.15)$$

и определяется только законом их распределения на промежутке $[0,T]$. Этого и следовало ожидать, исходя из простых вероятностных рассуждений.

Пример 3.2.1. Равномерное распределение моментов наблюдения на промежутке $[0,T]$ соответствует плотности

$$p(t) = \begin{cases} T^{-1}, & 0 \le t \le T, \\ 0, & t < 0,\ t > T. \end{cases} \qquad (3.2.16)$$

Подстановка (16) в (11) приводит к следующему выражению для характеристической функции:

$$H(\nu) = \frac{1}{i\cdot 2\pi\nu T}\left(1-e^{-i\cdot 2\pi\nu T}\right), \qquad (3.2.17)$$

так что из (14) следует:

$$R_N(\nu) = \frac{1}{N} + \left(1-\frac{1}{N}\right)\mathrm{sinc}^2(\nu T), \qquad (3.2.18)$$

где $\mathrm{sinc}(x) = \sin(\pi x)/(\pi x)$. При $N \gg 1$ формулы (18) и (7) принимают соответственно вид

$$R_N(\nu) \simeq \mathrm{sinc}^2(\nu T), \qquad (3.2.19)$$

$$\langle\langle P_N(\nu)\rangle\rangle \simeq \int_{-\infty}^{\infty} \mathrm{sinc}^2[(\nu-\nu')T]g(\nu')d\nu' \quad . \qquad (3.2.20)$$

Наконец, сравнивая последнее выражение с (1.5.33), мы видим, что

$$\langle\langle TP_N(\nu)\rangle\rangle \simeq g_T(\nu) \quad , \quad N \gg 1, \qquad (3.2.21)$$

– большое число «совершенно случайно» разбросанных по интервалу $[0,T]$ измерений эквивалентно непрерывному наблюдению случайного процесса $\xi(t)$ в течение промежутка времени T. Это утверждение, которое следует иметь в виду при планировании наблюдательных программ, высказывалось на основании интуитивных соображений Дамингом (1975).

Случайные отклонения от равномерности. Обратимся теперь к модели (1), когда исходная сетка моментов регистрации равномерна, но действительные моменты наблюдения процесса $\{t_k\}$ отклоняются от заданных случайным образом:

$$t_k = (k+\varepsilon_k)\Delta t, \quad k = 0, 1, \ldots, N-1. \qquad (3.2.22)$$

Поставленная задача может быть решена при достаточно общих предположениях о статистических свойствах совокупности $\{\varepsilon_k\}$, однако здесь мы ограничимся рассмотрением наиболее простого случая, позволяющего ясно увидеть суть тех изменений, которые вносят случайные отклонения от равномерности.

Итак, предположим, что $\{\varepsilon_k\}$ – попарно независимые случайные величины с одной и той же плотностью распределения вероятностей $p_\varepsilon(z)$, $-1/2 \leqslant z \leqslant 1/2$, и соответствующей характеристической функцией

$$H_\varepsilon(x) = \int_{-1/2}^{1/2} e^{-i\cdot 2\pi xz}\, p_\varepsilon(z)dz = \langle e^{-i\cdot 2\pi x\varepsilon_k}\rangle \qquad (3.2.23)$$

Подставляя (22) в определение (5) и учитывая формулу (23), получаем

$$R_N(\nu) = \frac{1}{N} + |H_\varepsilon(x)|^2 \frac{1}{N^2} \sum_{k\neq l} e^{-i\cdot 2\pi x(k-l)}, \qquad (3.2.24)$$

где обозначено $x = \nu\Delta t$. С другой стороны, плотность спектрального окна строго равномерного ряда $|W_N^{(e)}(\nu)|^2 = |\phi_N(x)|^2$ можно представить в виде

$$|\phi_N(x)|^2 = \frac{1}{N} + \frac{1}{N^2} \sum_{k\neq l} e^{-i\cdot 2\pi x(k-l)}. \qquad (3.2.25)$$

Из последних двух соотношений окончательно находим

$$R_N(x) = \frac{1}{N} + |H_\varepsilon(x)|^2 \left[|\phi_N(x)|^2 - \frac{1}{N} \right] . \qquad (3.2.26)$$

Поскольку $H_\varepsilon(0) = \phi_N(0) = 1$, равенство (6) выполнено.

Формула (26) совместно с основным соотношением (7) и определяет связь средней периодограммы и спектральной плотности процесса при случайных отклонениях от равномерности.

Как и следовало ожидать, строго равномерному ряду отвечает $R_N(\nu) = |\phi_N(x)|^2$, поскольку $p_\varepsilon(z) = \delta(z)$ и $H_\varepsilon(x) \equiv 1$. Ниже рассмотрены некоторые конкретные виды отклонений от равномерной сетки моментов наблюдения процесса.

Наибольший практический интерес представляет случай $N \gg 1$, когда можно полагать

$$R_N(\nu) \simeq |H_\varepsilon(x)\phi_N(x)|^2, \quad N \gg 1 , \qquad (3.2.27)$$

и число измерений N задает, как и прежде, только ширину максимумов функции $\phi_N(x)$. Именно формула (27) и будет использоваться в рассматриваемых далее примерах.

Пример 3.2.2. «Совершенно случайным» отклонениям от равномерности ряда соответствуют плотность

$$p_\varepsilon(z) = \begin{cases} 1, & -1/2 \le z \le 1/2 , \\ 0, & z < -1/2, \ z > 1/2 \end{cases} \qquad (3.2.28)$$

и характеристическая функция

$$H_\varepsilon(x) = \frac{\sin(\pi x)}{\pi x} \equiv \mathrm{sinc}(x) . \qquad (3.2.29)$$

Подставляя последнее выражение в (27), находим:

$$R_N(\nu) \simeq \left| \frac{\sin(\pi N x)}{\pi N x} \right|^2 = \mathrm{sinc}^2(Nx) , \qquad (3.2.30)$$

где $x = \nu\Delta t$ – безразмерная частота, а само соотношение (7) принимает вид

$$\langle P_N(\nu) \rangle \simeq \int_{-\infty}^{\infty} \mathrm{sinc}^2[(\nu-\nu')N\Delta t]\, g(\nu')\, d\nu' , \qquad (3.2.31)$$

фактически совпадающий с (20).

Как и следовало ожидать, закон (28) эквивалентен случаю равномерного случайного распределения моментов наблюдения на всем интервале $[0,T]$, который, в свою очередь, сводится при $N \gg 1$ к непрерывной регистрации процесса (пример 3.1.1).

Пример 3.2.3. Случайные малые отклонения от равномерной последовательности моментов наблюдения можно в случае их симметрии аппроксимировать нормальным законом, пренебрегая «крыльями» этого распределения, выходящими за промежуток $[-1/2, 1/2]$. Пусть σ_ε^2 – дисперсия отклонений ε_k; в соответствии со сказанным при $\sigma_\varepsilon^2 \ll 1$ положим

$$p_\varepsilon(z) \simeq \frac{1}{\sigma_\varepsilon\sqrt{2\pi}}\, e^{-z^2/(2\sigma_\varepsilon^2)}. \qquad (3.2.32)$$

Характеристическая функция нормального распределения равна

$$H_\varepsilon(x) = e^{-\frac{1}{2}(2\pi x\sigma_\varepsilon)^2}, \qquad (3.2.33)$$

так что

$$R_N(\nu) \simeq e^{-(2\pi x\sigma_\varepsilon)^2} \left|\phi_N(x)\right|^2, \quad x = \nu\Delta t. \qquad (3.2.34)$$

Мы сталкиваемся здесь с типичной для временных рядов, не слишком сильно отличающихся от равномерного, ситуацией: максимумы плотности спектра скважности на частотах $\nu_l = \pm l/\Delta t$, $l = 1, 2, \ldots$, соответствующих среднему шагу дискретизации Δt, последовательно убывают по мере удаления от центрального максимума $R_N(0) = 1$ (см., в частности, пример 1.6.2 и рис. 1.6.1). Следует подчеркнуть, что описываемое явление обусловлено именно непрерывностью набора длин интервалов между моментами наблюдения. Если, скажем, некоторые отсчеты строго равномерного ряда просто удалить, то для полученного неравномерного временного ряда максимумы $R_N(\nu)$ на частотах ν_l останутся равными 1. Качественно этот результат вполне понятен: гармоники, соответствующие ложным частотам (aliases), точно проходят через единичные отсчеты, расположенные по оси времени пусть и неравномерно, но через кратные некоторому фиксированному шагу Δt интервалы. Если же имеются отклонения от указанного распределения, то они сказываются на качестве аппроксимации синусоидой тем сильнее, чем выше ее частота (дефекты заметнее на высоких частотах). Поэтому скорость

убывания последовательных максимумов плотности спектрального окна может служить показателем степени неравномерности ряда в указанном смысле.

Что касается конкретного закона (34), то, например, отличие ближайших к центральному максимумов ($x = \pm 1/2$) от 1 составляет величину порядка $(\pi\sigma_\varepsilon)^2$.

§ 3.3. Отбеливание временных рядов

Термин *отбеливание* применяется для описания процесса выделения компонентов из временного ряда, при котором остаточный ряд близок по своим свойствам к белому шуму, т.е. уже не содержит информации.

Для наиболее простой модели «периодическая функция плюс шум» отбеливание сводится по возможности к более точному определению формы периодического колебания и дальнейшему вычитанию его (при условии сохранения фазы) из наблюдаемых отсчетов. Однако временной ряд, вообще говоря, не обязательно включает периодический компонент аддитивным образом, и тогда устранение последнего производится в соответствии с принятой моделью ряда. В § 1.1 мы обсуждали в качестве примера отбеливание ряда A; поскольку этот ряд имеет искусственное происхождение и, следовательно, характеристики случайного компонента полностью известны, мы могли констатировать близость остаточного ряда к исходной реализации шума (рис. 1.1.4 и 1.1.9).

Принятое выше определение предполагает, что процесс отбеливания включает в себя не только выделение периодических компонентов (линейчатого спектра), но и детерминированных или случайных компонентов с непрерывным спектром, заметно меняющимся в зависимости от частоты. Методы исследования случайного компонента ряда достаточно подробно рассматривались в предыдущей главе (см. в особенности § 2.6, посвященный моделям авторегрессии); процедура выделения детерминированного компонента - тренда - обсуждается особо в § 3.5. Поэтому в данном параграфе мы сосредоточим внимание на более узкой задаче отбеливания, заключающейся в выделении компонентов с дискретным спектром.

Следует отметить, что центрирование ря-

да, нередко выполняющееся в самом начале его исследования, можно рассматривать как отбеливание гармоники с нулевой частотой, преследующее цель устранить из периодограммы вместе с пиком на нулевой частоте и все связанные с ним ложные максимумы, обусловленные скважностью наблюдений. Такая операция, несомненно, полезна при поисках периодичности, однако в других случаях предварительное центрирование ряда может оказаться нецелесообразным, поскольку оно искажает истинный спектр.

Гармоническое отбеливание. Иногда априори известно, что периодическое колебание, если оно вообще присутствует, имеет чисто синусоидальную форму. Так, например, обстоит дело при изучении переменности блеска звезды, обусловленной ее вращением и неоднородной структурой поверхности (при этом характерное время изменчивости пятен должно существенно превосходить период вращения звезды). И в более сложных ситуациях гармоническое отбеливание применяется настолько часто, что имеет смысл рассмотреть его отдельно.

Предположим, что процесс описывается аддитивной моделью

$$y(t) = A_0 + A_1\cos(2\pi\nu_1 t - \varphi_1) + \xi(t) = \\ = A_0 + a_1\cos(2\pi\nu_1 t) + b_1\sin(2\pi\nu_1 t) + \xi(t), \qquad (3.3.1)$$

где коэффициенты $a_1 = A_1\cos\varphi_1$, $b_1 = A_1\sin\varphi_1$, а $\xi(t)$ - белый шум с нулевым средним и (неизвестной) дисперсией σ^2. Наблюдениям в моменты $\{t_k\}$, $k = 0, 1, \ldots, N-1$ отвечает совокупность отсчетов $\{y_k\}$, наилучшую аппроксимацию которой естественно искать в виде

$$g(t) = \hat{A}_0 + \hat{a}_1\cos(2\pi\hat{\nu}_1 t) + \hat{b}_1\sin(2\pi\hat{\nu}_1 t), \qquad (3.3.2)$$

где частота $\hat{\nu}_1$ обычно соответствует положению максимума периодограммы, а коэффициенты $\hat{A}_0$, $\hat{a}_1$ и $\hat{b}_1$ определяются из условия наилучшего в среднеквадратическом смысле согласия между $\{y_k\}$ и $\{g_k\}$. Указанное условие приводит к системе линейных уравнений

$$\hat{a}_1\sum\cos\vartheta_k+\hat{b}_1\sum\sin\vartheta_k+\hat{A}_0N=\sum y_k,$$

$$\hat{a}_1\sum\cos^2\vartheta_k+\hat{b}_1\sum\sin\vartheta_k\cos\vartheta_k+\hat{A}_0\sum\cos\vartheta_k=\sum y_k\cos\vartheta_k, \qquad (3.3.3)$$

$$\hat{a}_1\sum\sin\vartheta_k\cos\vartheta_k+\hat{b}_1\sum\sin^2\vartheta_k+\hat{A}_0\sum\sin\vartheta_k=\sum y_k\sin\vartheta_k,$$

где $\vartheta_k=2\pi\hat{\nu}_1 t_k$. После нахождения $\hat{a}_1$ и $\hat{b}_1$ можно определить оценки амплитуды $\hat{A}_1=(\hat{a}_1^2+\hat{b}_1^2)^{1/2}$ и начальной фазы колебания

$$\hat{\varphi}_1=\begin{cases}\operatorname{arctg}\hat{b}_1/\hat{a}_1, & \hat{a}_1\geqslant 0,\ \hat{b}_1\geqslant 0,\\ \operatorname{arctg}\hat{b}_1/\hat{a}_1+\pi, & \hat{a}_1<0,\\ \operatorname{arctg}\hat{b}_1/\hat{a}_1+2\pi, & \hat{a}_1>0,\ \hat{b}_1<0,\end{cases} \qquad (3.3.4)$$

задающей угловое расстояние между началом отсчета времени $t=0$ и ближайшим к нему максимумом синусоиды ($0\leqslant\hat{\varphi}_1\leqslant 2\pi$). Дисперсия остаточного ряда $\{y_k-g_k\}$ может служить оценкой дисперсии σ^2 шума $\xi(t)$. Далее возможен итеративный процесс, включающий последовательное уточнение параметров колебания.

Полигармоническое отбеливание. На практике априорная информация о форме искомого периодического колебания чаще всего отсутствует, и тогда мы должны предполагать для аддитивного процесса

$$y(t)=A_0+\sum_{n=1}^{\infty}\left[a_n\cos(2\pi n\nu_0 t)+b_n\sin(2\pi n\nu_0 t)\right]+\xi(t), \qquad (3.3.5)$$

где $\nu_0^{-1}\equiv P_0$ - период колебания (мы отождествляем обозначения ν_0 и ν_1 для основной частоты). Предположим далее, что амплитуда первой гармоники $A_1=(a_1^2+b_1^2)^{1/2}$ максимальна (т.е. форма колебания не слишком сильно отличается от синусоидальной), отношение сигнала к шуму A_1^2/σ^2 достаточно велико, а распределение моментов наблюдения $t_0,\ldots,t_{N-1}$ не содержит пекулярностей, приводящих к флуктуациям периодограммы на частотах высших гармоник $2\nu_0$, $3\nu_0,\ldots$ При этих условиях, обычно выполняющихся в реальных задачах, наибольший максимум периодограммы ряда $\{y_k\}$ находится на частоте $\hat{\nu}_0$, близкой к теоретической частоте ν_0. Описанное в § 1.1

сворачивание ряда с периодом $\hat{P}_0 = \hat{\nu}_0^{-1}$ приводит к фазовой диаграмме типа изображенной на рис. 1.1.3. Если бы шум $\xi(t)$ был пренебрежимо мал, моменты наблюдения $\{t_k\}$ соответствовали представительному набору фаз, а частоты ν_0 и $\hat{\nu}_0$ отличались незначительно, то распределение точек на фазовой диаграмме достаточно близко обрисовало бы форму искомого периодического колебания, как это и показывает рис. 1.1.8. В действительности фазовая диаграмма очень редко бывает столь хороша, она чаще напоминает диаграмму, представленную на рис. 1.1.3, и для определения формы колебания следует каким-либо образом аппроксимировать распределение точек плавной кривой.

Выбор способа аппроксимации, разумеется, произволен. Иногда используются полиномиальное или сплайновое приближения, однако удобнее, по-видимому, аппроксимация тригонометрическим полиномом

$$g_m(x) = \hat{A}_0 + \sum_{n=1}^{m}\left[\hat{a}_n \cos(2\pi n x) + \hat{b}_n \sin(2\pi n x)\right] , \qquad (3.3.6)$$

непрерывным вместе со своими производными при идентичных по смыслу фазах $x = 0$ и $x = 1$ (определение фазы дается формулой (1.1.18)). Надлежащий выбор порядка m (см. ниже) дает оценки коэффициентов $\hat{A}_0$, $\hat{a}_n$ и $\hat{b}_n$, близкие к соответствующим теоретическим значениям в (5), а дисперсия остаточного ряда

$$\xi_m(t_k) \equiv y(t_k) - g_m[x(t_k)] \qquad (3.3.7)$$

может рассматриваться в качестве оценки дисперсии шума $\xi(t)$

Нахождение коэффициентов аппроксимирующей функции $g_m(x)$ основывается чаще всего на методе наименьших квадратов, исходящем из условия

$$\sum_{k=0}^{N-1} \xi_m^2(t_k) = \min . \qquad (3.3.8)$$

Для произвольного заданного значения m это условие сводится к следующей системе $2m+1$ линейных алгебраических уравнений,

определяющих коэффициенты $\hat{A}_0$, $\hat{a}_n$, $\hat{b}_n$, $n = 1, \ldots, m$:

$$\begin{aligned}
&\hat{A}_0 + \sum_{l=1}^{m} \hat{a}_l c_{l0} + \sum_{l=1}^{m} \hat{b}_l r_{l0} = \bar{y},\\
&\hat{A}_0 c_{0n} + \sum_{l=1}^{m} \hat{a}_l c_{ln} + \sum_{l=1}^{m} \hat{b}_l r_{ln} = U_n,\\
&\hat{A}_0 r_{n0} + \sum_{l=1}^{m} \hat{a}_l r_{nl} + \sum_{l=1}^{m} \hat{b}_l s_{ln} = V_n,
\end{aligned} \qquad (3.3.9)$$

где

$$\begin{aligned}
&\bar{y} = \frac{1}{N} \sum_{0}^{N-1} y_k \ , \quad \vartheta_k = 2\pi x_k,\\
&c_{ln} = c_{nl} = \frac{1}{N} \sum_{0}^{N-1} \cos(l\vartheta_k)\cos(n\vartheta_k),\\
&r_{ln} = \frac{1}{N} \sum_{0}^{N-1} \sin(l\vartheta_k)\cos(n\vartheta_k),\\
&s_{ln} = s_{nl} = \frac{1}{N} \sum_{0}^{N-1} \sin(l\vartheta_k)\sin(n\vartheta_k),\\
&U_n = \frac{1}{N} \sum_{0}^{N-1} y_k \cos(n\vartheta_k),\\
&V_n = \frac{1}{N} \sum_{0}^{N-1} y_k \sin(n\vartheta_k).
\end{aligned} \qquad (3.3.10)$$

Амплитуды и начальные фазы отдельных гармоник определяются по формулам, аналогичным приведенным в предыдущем пункте. Оценка дисперсии остаточного ряда (7) равна

$$\Delta_m^2 = \frac{1}{N-1} \sum_{k=0}^{N-1} \xi_m^2(t_k) \ . \qquad (3.3.11)$$

Мы не имеем здесь возможности обсуждать все случаи, когда не выполняются принятые ранее условия, скажем, максимальна амплитуда одной из высших гармоник или случайной периодической последовательности, возникшей в шуме. Эти случаи нетрудно рассмотреть на основании соображений, аналогичных высказанным.

Отбеливание основной периодичности, содержащейся в рассматриваемом временном ряду, делает более контрастными возможные периодичности меньшей амплитуды, которые, в свою оче-

редь, можно подвергнуть процедуре последовательного отбеливания. Если нет оснований отличать остаточный ряд на какомлибо шагу от шума (соответствующие критерии обсуждаются ниже), то мы приходим в итоге к модели временного процесса, включающей конечное число периодических компонентов и шум.

Заметим, наконец, что возможно одновременное отбеливание нескольких периодических компонентов с оптимальным в смысле метода наименьших квадратов подбором коэффициентов аппроксимирующей функции (см. Пельт, 1980). Однако эту процедуру целесообразно выполнять лишь на заключительном этапе исследования, когда структура временного ряда уже выяснена при последовательном отбеливании отдельных составляющих.

Особенности отбеливания неравномерных рядов. При фиксированном интервале Δt между моментами наблюдения плотность спектрального окна $|W_N(\nu)|^2$ имеет вид единичной «гребенки» (см. рис. 2.1.1), отдельные «зубцы» которой расположены на частотах, кратных частоте выборки $\nu_s = \Delta t^{-1}$. Если число отсчетов достаточно велико, то функцию $|W_N(\nu)|^2$ можно считать равной нулю в главном частотном интервале $0 \leqslant \nu \leqslant \nu_c = (2\Delta t)^{-1}$, и потому отбеливание гармонического колебания некоторой частоты ν_0 мало влияет на вид спектральной плотности на других частотах, если только эти последние не совпадают с последовательностью сопряженных ν_0 частот.

С иным положением мы сталкиваемся при отбеливании неравномерных рядов. На рис. 3.3.1 приведены плотность спектрального окна $|W_N(\nu)|^2$ и периодограмма $P_N(\nu)$ синусоиды $f(t) = A\cos[2\pi\nu_0(t-t_0)]$ при $A = 0{,}31$, $\nu_0 = 0{,}3151$ цикл/сут, $N = 49$ и совокупности $\{t_k\}$, совпадающей с действительным распределением моментов наблюдения цефеиды BK Cen (Леотта-Жанин, 1967; табл. П.7.4). Как видно, $|W_N(\nu)|^2$ имеет заметные максимумы не только на частотах, кратных 1 цикл/сут, так что когда спектр скважности «цепляется» по терминологии § 1.6 за пики на частотах $\pm\nu_0$, в наблюдаемой периодограмме возникает целый набор максимумов - «лес» - сравнительно небольшой амплитуды, не имеющий отношения к ошибкам измерения и вообще к шумовому компоненту. Очевидно, что отбеливание колебания с частотой ν_0 уничтожит не только пик на этой частоте и сопряженных ей частотах, но и весь «лес».

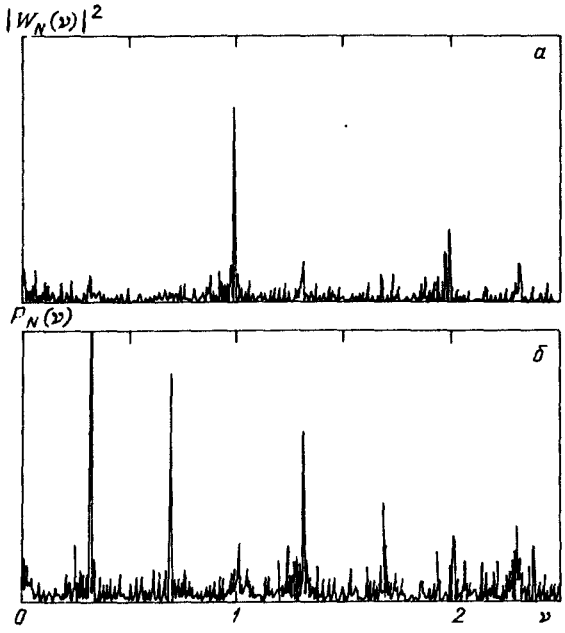


Рис. 3.3.1. Плотность спектрального окна для моментов наблюдения звезды BK Cen (а) и периодограмма модельного синусоидального ряда с частотой 0,3151 цикл/с для этих моментов времени

Другая особенность отбеливания неравномерных рядов связана с явлением подмены частот. Неравномерность распределения моментов наблюдения приводит к убыванию амплитуд последовательных пиков спектра скважности на частотах, кратных основной частоте модуляции ν_w (см. рис. 1.6.1, где ν_w отвечает суточной модуляции). Если частота периодического компонента ν_0 меньше $\nu_w/2$, то на сопряженных ν_0 частотах будут наблюдаться пики, почти всегда имеющие меньшую амплитуду по сравнению с основным пиком (рис. 1.6.1), что соответствует убыванию максимумов $|W_N(\nu)|^2$. Ясно, что отбеливание основного, наиболее выделяющегося колебания с частотой ν_0 уничтожит и все сопряженные максимумы. Если же отбелить колебание на одной из сопряженных частот, то дисперсия остаточного ряда Δ_m^2 окажется не столь малой, как в первом случае. Однако периодограмма неравномерного ряда может иметь заметную положительную флуктуацию на одной из сопря-

женных частот, например на ν_w-ν_0, которая отражает случайную периодичность, обусловленную распределением моментов наблюдения. Тогда наложение сопряженного максимума и флуктуации может привести к тому, что основной пик на частоте ν_0 окажется не самым выделяющимся в спектре, а колебание сопряженной частоты будет ошибочно принято за основное.

Оптимальный порядок аппроксимирующего полинома. При обсуждении отбеливания ряда А в § 1.1 мы произвольно выбрали значение $m = 4$ для порядка аппроксимирующего полинома $g_m(x)$. Получившееся в результате приближенное описание распределения точек на фазовой диаграмме иллюстрируется рис. 1.1.3. Мы видим, что $g_4(x)$ правильно передает основные черты «кривой блеска», в частности, ее несимметричность (ср. с истинной кривой на рис. 1.1.8). По-видимому, в данном случае не следует ограничиваться синусоидальным отбеливанием ($m = 1$), однако и выбор $m = 4$ не представляется оптимальным, так как ощущается излишняя «свобода» $g_4(x)$, приводящая к отбеливанию и случайных флуктуаций.

Итак, мы пришли к проблеме выбора оптимального порядка тригонометрического полинома, аппроксимирующего временной ряд. Само существование оптимального значения m определяется наличием шума, так как описание чисто детерминированного периодического сигнала становится, вообще говоря, все более точным по мере возрастания m. Формально с ростом m уточняется и описание реального ряда, включающего случайный компонент; в частности, последовательно уменьшается величина Δ_m^2, определенная выражением (11). Ясно, однако, что чрезмерно высокий порядок полинома не отражает существа дела: в рассмотрение начинают входить случайные периодичности, неизбежно присутствующие в шуме. По этой причине уменьшение Δ_m^2 с ростом m не обязательно свидетельствует об улучшении качества аппроксимации искомой периодической зависимости.

Исходным пунктом для выбора критерия оптимальности отбеливания является следующее соображение: каждый раз при увеличении порядка аппроксимации вновь вычисленные коэффициенты основываются на *тех же значениях* временного ряда, т.е. увеличивается и число связей, накладываемых на выборку. Полное число связей равно $2m+1$ - числу коэффициентов,

входящих в аналитическое представление (6). В соответствии с этим качество аппроксимации характеризуется величиной

$$L_m = \frac{1}{N-2m-1} \sum_{k=0}^{N-1} \xi_m^2(t_k), \qquad (3.3.12)$$

где $\xi_m(t_k)$ - заданный соотношением (7) остаточный ряд (Пустыльник, 1968, п.9.2). В отличие от (11) здесь уменьшение суммы квадратов остатков компенсируется ростом числа связей, так что L_m может достигать минимума при некотором значении m_0, которое и следует считать оптимальным. Вывод (12) на основании качественных соображений приводится ниже.

Рассмотрим с точки зрения изложенного подхода отбеливание периодической зависимости с частотой $\hat{\nu}_0 = 0,3705$ в ряде А. На рис. 3.3.2 представлены зависимости параметров Δ_m^2 и L_m от порядка аппроксимации m. Очевидно, при $m = 0$ мы имеем $L_0 = \Delta_0^2$ - дисперсию исходного ряда. Как уже отмечалось, остаточная дисперсия Δ_m^2 уменьшается с ростом m. В то же время параметр качества аппроксимации L_m достигает первого минимума при $m = 2$, а далее вплоть до $m = 5$ медленно возрастает. Если бы при больших значениях m расчеты не производились, мы выбрали бы в качестве оптимального значения $m_0 = 2$, однако существует глубокий минимум при $m = 6$, и это заставляет более внимательно проанализировать процесс последовательного отбеливания. Шестая гармоника основной частоты 0,3705 имеет частоту 2,2230; ввиду явления наложения частот она неразличима с частотой 0,2230, на которой случайно оказался заметный пик периодограммы $\nu' = 0,228$ (см. рис. 1.1.2), обусловленный флуктуацией шума. Таким образом, вторичный минимум L_m при $m = 6$ связан лишь с тем, что аппроксимирующий полином $g_6(x)$ принимает случайный пик за шестую гармонику и устраняет его при отбеливании. Поскольку максимум на частоте $\nu' = 0,228$ нельзя считать значимым (см. соответствующую оценку в примере 2.3.1), не следует придавать значения и вторичному минимуму L_m, а в качестве оптимального порядка $g_m(x)$ целесообразно выбрать $m_0 = 2$.

Последние замечания относительно случайных максимумов периодограммы показывают, что не всякое изменение L_m статистически значимо. Так, при переходе от $m = 2$ к $m = 3$ для

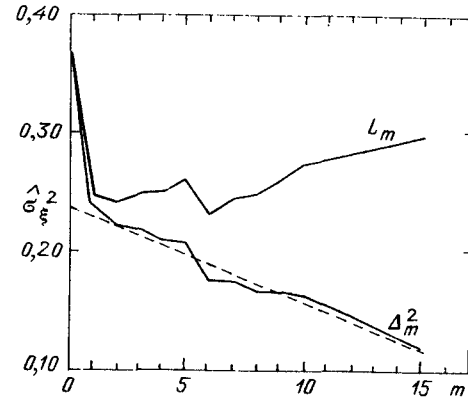


Рис. 3.3.2. Дисперсия остаточного ряда Δ^2_m и параметр качества аппроксимации L_m в зависимости от порядка аппроксимирующего полинома m для ряда *A*

ряда A параметр L_m изменяется всего на 2,4 %, что вряд ли заставит нас отдать явное предпочтение отбеливанию именно второго порядка. Величины L_m и L_{m+1} можно рассматривать как дисперсии двух выборок, и вопрос о значимости изменения параметра качества решить, как это обычно делается для дисперсий, на основе *статистики Фишера* $F_m \equiv L_m/L_{m+1}$ с числом степеней свободы $f_1 = N-2m-1$ и $f_2 = N-2m-3$ (см. Пустыльник, 1968, п.9.3).

Заметим, что не всегда параметр L_m достигает минимального значения при разумных m; тогда, по-видимому, следует прекратить повышение порядка аппроксимации при тех значениях m, для которых изменение F_m уже становится статистически незначимым.

Наряду с изучением зависимости L_m от порядка аппроксимации полезно следить за изменениями вида фазовой диаграммы и периодограммы остаточного ряда при росте m. Соответствующие иллюстрации для ряда A при m = 1, 2, 6 и 15 представлены на рис. 3.3.3 и 3.3.4. Следует обратить внимание на появление «ряби» при m = 6, а также на чрезмерную свободу полинома при m = 15, которая позволяет отслеживать многие быстрые флуктуации шума.

На рис. 3.3.2 привлекает внимание тот факт, что в пределах статистических флуктуаций изменение остаточной дисперсии Δ^2_m происходит при $m \geq 2$ практически *линейно* с ростом m. Нетрудно понять, что такого изменения Δ^2_m и следовало ожидать на основании самых общих соображений. Действительно,

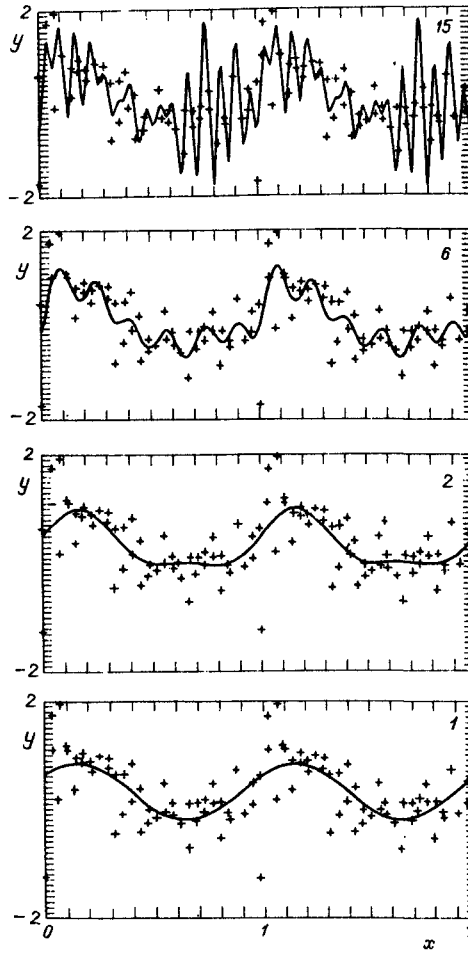


Рис. 3.3.3. Аппроксимация фазовой диаграммы ряда А тригонометрическими полиномами степени m = 1, 2, 6 и 15

формула (2.3.5) в случае белого шума $\xi(t)$, т.е. при $g_\xi(\nu)$ = const, дает $\sigma_\xi^2 = g_\xi\, 2\nu_c = g_\xi\, (\Delta t)^{-1}$. С другой стороны, центрирование ряда устраняет из спектральной плотности полосу шириной $\delta\nu \simeq T^{-1}$ вблизи $\nu = 0$, а каждое гармоническое отбеливание с частотой ν_k устраняет такие же полосы около частот $\pm\nu_k$. Поэтому после отбеливания белого шума полиномом порядка m дисперсия остаточного ряда равна приближенно

$$\frac{1}{N}\sum_{k=0}^{N-1}\xi_m^2(t_k) \simeq \sigma_\xi^2 - \frac{g_\xi}{T} - g_\xi\frac{2m}{T} =$$

$$= \frac{g_\xi}{\Delta t}\left(1 - \frac{2m+1}{T/\Delta t}\right) = \sigma_\xi^2\left(1 - \frac{2m+1}{N}\right), \qquad (3.3.13)$$

откуда

$$\Delta_m^2 = \sigma_\xi^2\left(1 - \frac{2m}{N-1}\right). \qquad (3.3.14)$$

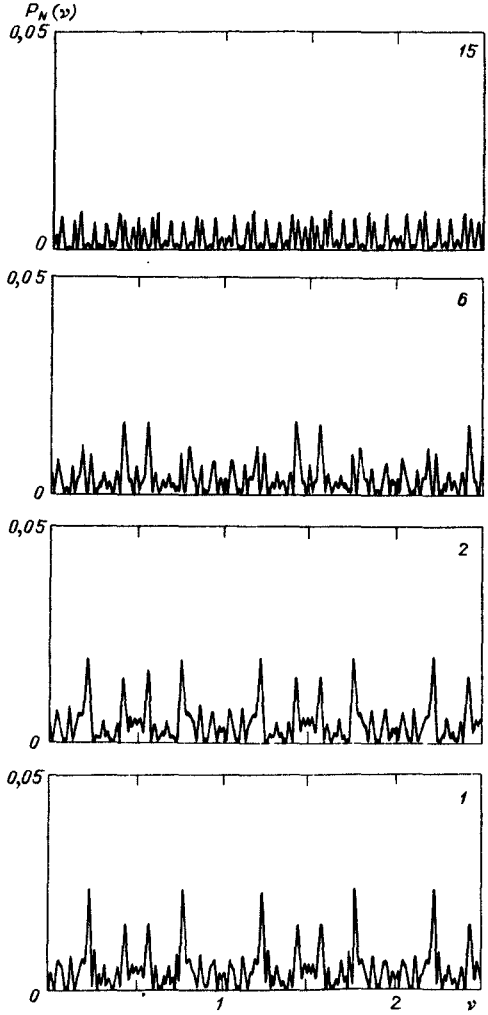


Рис. 3.3.4. Периодограммы остаточного ряда после отбеливания полиномами степени m = 1, 2, 6 и 15

Таким образом, величина Δ_m^2 убывает линейно с ростом m при отбеливании шума с постоянной спектральной плотностью. Важное значение этого результата заключается в том, что на его основе может быть построен критерий для проверки случайности остаточного ряда и, следовательно, остановки процесса аппроксимации: как только зависимость Δ_m^2 от m становится приблизительно линейной, можно считать, что остаточный ряд уже близок к чисто случайному.

Что касается непосредственно ряда A, то аппроксимация при помощи метода наименьших квадратов кривой Δ_m^2 на рис. 3.3.2 в пределах $2 \leqslant m \leqslant 15$ дает

$$\Delta_m^2 \simeq 0,238(1-0,344m). \qquad (3.3.15)$$

При $m = 0$ мы находим отсюда оценку дисперсии шума $\hat{\sigma}_\xi^2 = 0,238$, в то время как действительное значение дисперсии той реализации шума, которая входит в ряд A, равно 0,230 (дисперсия в теоретической модели шума равна 1/4, см. § 1.1). Заметим также, что и угловые коэффициенты в (14) и (15) приблизительно согласуются между собой при $N = 50$. Оценив на основании (14) дисперсию шума $\hat{\sigma}_\xi^2$, мы получаем возможность прекратить повышение порядка аппроксимации, как только Δ_m^2 уменьшится до $\hat{\sigma}_\xi^2$. В частности, для ряда A находим $m_0 \simeq 1 \div 2$ в согласии с нашими более ранними оценками.

Казалось бы, изложенный метод выбора оптимального значения m_0 не связан с методом, опирающимся на критерий качества аппроксимации L_m, однако в действительности эти методы эквивалентны. В самом деле, из (11) и (14) следует оценка дисперсии шума

$$\hat{\sigma}_\xi^2 = \frac{1}{N-2m-1} \sum_{k=0}^{N-1} \xi_m^2(t_k) , \qquad (3.3.16)$$

совпадающая с введенным согласно (12) параметром L_m. Этот результат поясняет смысл параметра качества аппроксимации и

позволяет дать более ясную трактовку процедуры выбора оптимального значения m_0. Именно, процесс увеличения порядка m следует остановить, как только оценка дисперсии шума достигнет статистически значимого минимума.

При выборе m_0 можно опираться и на другие критерии, использующие отличные от (12) характеристики качества. Вместе с тем проведенный выше анализ свидетельствует о том, что здесь вряд ли возможно ограничиться каким-либо одним формальным критерием, позволяющим обойтись без участия исследователя (см., например, обсуждение минимума $m = 6$ для ряда А). Напомним также, что и при выборе порядка авторегрессионной модели формальный критерий Акаике не всегда эффективен. Поэтому представляется уместным сформулировать некоторые необходимые правила, которых следует придерживаться при отбеливании временного ряда:

1. Не ограничиваться каким-либо одним значением m, но провести последовательно отбеливание разных порядков, сравнивая результаты на каждом шагу.

2. Проверить устойчивость выводов относительно выбора m_0.

3. Дать однозначное описание всех использованных при анализе процедур.

Пример 3.3.1. Цефеида BK Cen относится к немногочисленному классу звезд типа TU Cas, кривые блеска которых обнаруживают присутствие двух независимых периодических колебаний. Значения периодов колебаний находятся обычно в диапазоне от 1 до 7 сут, а их отношения сосредоточены в узком интервале $0{,}696 < P_1/P_0 < 0{,}711$ (Стоби, 1977).

Согласно анализу Леотта-Жанин (1967) фотоэлектрических наблюдений BK Cen, выполненных ван Хаутеном (табл. П.7.4, T= = 136,76 суток, N = 49), для этой звезды P_0= 3,17389 сут, P_1 = 2,2366 сут (соответствующие частоты ν_0=0,3151 цикл/сут, ν_1=0,4471 цикл/сут) и отношение P_1/P_0=0,705. Спектр скваж-

ности указанного ряда наблюдений был представлен на рис. 3.3.1; периодограмма, фазовая диаграмма для частоты ν_0 и результат отбеливания основного колебавия с частотой ν_0 при $m = 3$ приведены на рис. 3.3.5.

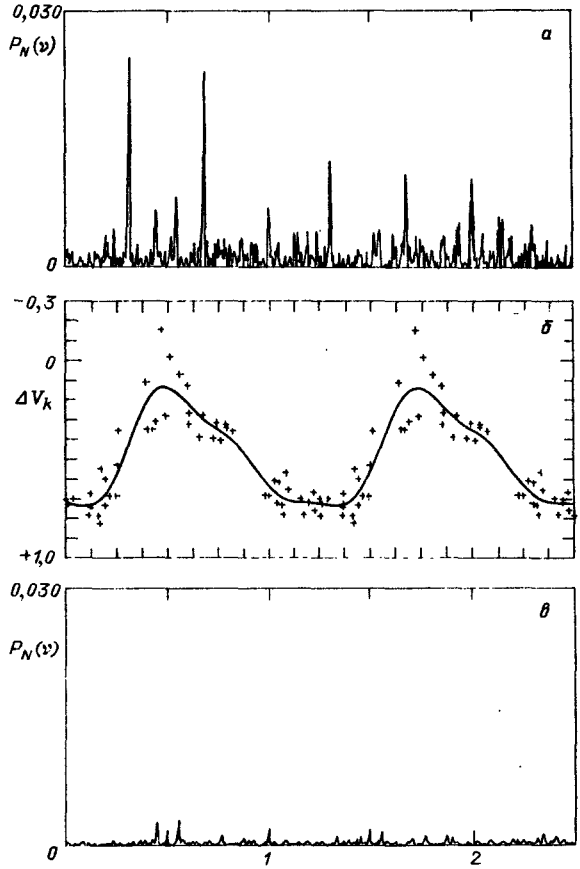


Рис. 3.3.5. Периодограмма колебаний блеска двухмодовой цефеиды BK Cen (а), фазовая диаграмма для $\nu_0 = 0,3151$ цикл/с (б) и результат отбеливания колебания с частотой ν_0 при $m = 3$ (в)

Периодограмма BK Cen действительно обнаруживает максимумы на частотах ν_0 и ν_1, отвечающих двум модам колебаний. Вместе с тем привлекают внимание некоторые особенности, требующие специального дальнейшего изучения: 1) максимум на частоте $1-\nu_0 \approx 0,6849$ цикл/сут, сопряженный основному, характеризуется большей интенсивностью, чем этого следовало ожидать на основании спектра скважности; 2) сопряженный вто-

ричному максимум на частоте $1-\nu_1 \approx 0{,}5529$ цикл/сут вообще выше вторичного максимума; 3) отбеливание основного колебания с частотой ν_0 сильно редуцирует не только соответствующие ложные максимумы и «лес» скважности, но и максимумы на частотах ν_1, $1-\nu_1$, отвечающие вторичному колебанию. Более того, если в исходных данных отбелить только колебание с частотой 0,685 цикл/сут, вторичные максимумы редуцируются еще сильнее. К сожалению, ВК Cen недоступна для наблюдений в северном полушарии.

Данный пример наглядно иллюстрирует взаимосвязь различных частей периодограммы неравномерного временного ряда – отбеливание лишь на одной (скажем, низкой) частоте существенным образом меняет вид спектральной плотности во всем диапазоне вплоть до весьма высоких частот. Отсюда, в частности, видно, насколько опасно при анализе ограничиваться, как это часто делают, узким спектральным участком и сразу переходить к оценке значимости расположенных в нем максимумов.

Критерии случайности. Остающийся после удаления тренда и отбеливания гармонических составляющих ряд предполагается выборкой из некоторого случайного процесса, не обязательно представляющего собой белый шум. Если спектральная плотность ряда значимо отличается от постоянной (соответствующие критерии обсуждаются, например, Дженкинсом и Ваттсом, 1968), то дальнейший анализ может включать в себя подбор авторегрессионной модели или различных ее обобщений (§ 2.6). В наиболее простом случае, часто тем не менее встречающемся на практике, у исследователя нет оснований предполагать реальность сложного внутреннего механизма, управляющего рядом, и он хотел бы проверить возможность того, что остаточный ряд является просто набором независимых друг от друга отсчетов, т.е. белым шумом. Как отмечают Кендалл и Стюарт (1966), «нет предела числу критериев, которые можно построить для этой цели». Ниже мы приводим два простейших критерия случайности ряда, не зависящих от функции распределения его отсчетов. Дальнейшие сведения по обсуждаемому воп-

росу можно найти во многих руководствах, и в частности, цитированных выше. Помимо этого достаточно удобным критерием белого шума может служить обсуждавшаяся выше линейность зависимости Δ_m^2 от порядка аппроксимации.

1) Ч и с л о т о ч е к п о в о р о т а . Критерий заключается в подсчете числа «пиков» и «впадин» ряда, где первые определяются как значения, превосходящие соседние значения, а вторые - как значения, меньшие обоих соседних.

Пусть ряд длины N состоит из независимых друг от друга отсчетов с некоторой функцией распределения $F(x)$. Вероятность того, что произвольно взятый внутренний отсчет ряда заключен в интервале $[x, x+dx]$, есть $dF(x)$. Условная вероятность того, что при этом соседние отсчеты находятся либо ниже, либо выше взятого, равна $F^2(x)+[1-F(x)]^2 = 1-2F(x)[1-F(x)]$. Следовательно, вероятность того, что взятый отсчет является поворотной точкой, равна

$$\int_0^1 [1-2F(1-F)]dF = \frac{2}{3} \qquad (3.3.17)$$

независимо от вида $F(x)$. Ряд включает $N-2$ внутренние точки, так что среднее значение μ числа точек поворота совершенно случайного ряда определяется формулой

$$\bar{\mu} = \frac{2}{3}(N-2). \qquad (3.3.18)$$

Несколько более сложно находится дисперсия числа точек поворота (см. Кендалл, 1976):

$$\sigma_\mu^2 = \frac{1}{90}(16N-29). \qquad (3.3.19)$$

Например, для ряда длины $N = 50$ ожидается $\bar{\mu} = 32$ точки поворота при стандартном отклонении $\sigma_\mu = 2,93$.

Поскольку распределение μ довольно быстро стремится к нормальному распределению, применение описываемого критерия не составляет труда.

2) Ч и с л о п о л о ж и т е л ь н ы х р а з н о с т е й последовательных значений ряда с гакже характеризует степень его случайности. Кендалл (1976) приводит значения моментов

$$\bar{c} = \frac{1}{2}\,(N-1),$$
$$\sigma_c^2 = \frac{1}{12}\,(N+1), \qquad (3.3.20)$$

однако замечает, что критерий скорее подходит для проверки наличия линейного тренда, поскольку при колебательном движении число точек возрастания ряда всегда будет приблизительно равно $N/2$.

§ 3.4. Алгоритм CLEAN

В § 1.6 было показано, что истинный спектр детерминированного процесса

$$F(\nu) = \int_{-\infty}^{\infty} e^{-i\cdot 2\pi\nu t} f(t)dt\ , \qquad (3.4.1)$$

его дискретный спектр

$$D_N(\nu) = \frac{1}{N}\sum_{k=0}^{N-1} e^{-i\cdot 2\pi\nu t_k} f(t_k) \qquad (3.4.2)$$

и спектральное окно данных

$$W_N(\nu) = \frac{1}{N}\sum_{k=0}^{N-1} e^{-i\cdot 2\pi\nu t_k} \qquad (3.4.3)$$

связаны интегральным уравнением

$$D_N(\nu) = \int_{-\infty}^{\infty} W_N(\nu-\nu')F(\nu')d\nu'. \qquad (3.4.4)$$

Функции $D_N(\nu)$ и $W_N(\nu)$ полностью определены совокупностью моментов наблюдения и отсчетов временного ряда $\{t_k,\ f_k\}$, т.е. считаются заданными. При наличии шума соотношения между искомыми и наблюдаемыми величинами усложняются (см. § 1.6); для ясности изложения мы ограничимся, следуя работе Робертса и др. (1987), рассмотрением детерминированных процессов.

В общем случае дискретный спектр существенно отличается от истинного спектра, поскольку первый содержит ложные пики, обусловленные конкретным распределением $\{t_k\}$. Как уже отмечалось в § 1.6, непосредственное решение (4) для произвольного вида $F(\nu)$ стандартными методами обращения уравне-

ний типа свертки неэффективно. Однако для *гармонического* процесса $f(t) = A\cos(2\pi\nu_0 t - \varphi_0)$ уравнение (4) легко позволяет найти амплитуду A и начальную фазу φ_0. В самом деле, положим

$$a \equiv \frac{A}{2} e^{-i\varphi_0} \qquad (3.4.5)$$

и перепишем соотношения (1.4.14) и (1.6.25), определяющие истинный и дискретный спектры $f(t)$, соответственно в виде

$$F(\nu) = a\delta(\nu-\nu_0) + a^*\delta(\nu+\nu_0) \ , \qquad (3.4.6)$$

$$D_N(\nu) = aW_N(\nu-\nu_0) + a^*W_N(\nu+\nu_0) \ . \qquad (3.4.7)$$

На каждой частоте дискретный спектр содержит вклад функций спектрального окна, центрированных на частотах максимумов истинного спектра $\pm\nu_0$ (как говорят, спектр «загрязнен» ложной периодичностью). В частности, на самой частоте ν_0 имеем:

$$D_N(\nu_0) = a + a^*W_N(2\nu_0) \ , \qquad (3.4.8)$$

где мы воспользовались (1.6.22). Запишем решение (8) относительно a в виде:

$$a = \frac{D_N(\nu_0) - D_N^*(\nu_0)W_N(2\nu_0)}{1 - |W_N(2\nu_0)|^2} \ . \qquad (3.4.9)$$

Эта формула и определяет комплексную амплитуду гармонического колебания через наблюдаемые величины.

На практике ввиду искажения формы линий истинного спектра точное значение частоты ν_0 неизвестно, однако можно полагать, что оно находится в пределах интервала разрешения $\delta\nu \simeq T^{-1}$ около частоты максимума $\hat{\nu}_0$ наблюдаемого спектра. Поэтому оценку параметра a на основании дискретного спектра $D_N(\nu)$ можно представить следующим образом:

$$a(D_N, \hat{\nu}_0) = \frac{D_N(\hat{\nu}_0) - D_N^*(\hat{\nu}_0)W_N(2\hat{\nu}_0)}{1 - |W_N(2\hat{\nu}_0)|^2} \ . \qquad (3.4.10)$$

Вычислив $a(D_N, \hat{\nu}_0)$, можно устранить из наблюдаемого спектра $D_N(\nu)$ весь вклад $a(D_N, \hat{\nu}_0)W_N(\nu-\hat{\nu}_0) + a^*(D_N, \hat{\nu}_0)W_N(\nu+\hat{\nu}_0)$, обус-

ловленный компонентом с частотой $\hat{\nu}_0$ (при этом исчезают и связанные с рассматриваемым компонентом ложные пики). Если остаточный спектр $R^{(1)}(\nu)$ не содержит выделяющихся максимумов, искомое разложение процесса считается законченным; в противном случае описанная процедура повторяется для частот максимумов остаточного спектра. Конечным результатом последовательности K операций «очищения» спектра является совокупность комплексных амплитуд $\{a_i\}$ и остаточный спектр $R^{(K)}(\nu)$.

Изложенная процедура представляет собой одномерный вариант алгоритма CLEAN, который был разработан Хёгбомом (1974) для анализа двумерных распределений яркости в радиоастрономии, построенных при помощи апертурного синтеза.

Практическое использование алгоритма CLEAN в анализе временных рядов производится следующим образом (Робертс и др., 1987).

Обозначим $R^{(0)}(\nu) \equiv D_N(\nu)$.

1) Начиная с $i = 1$, найдем частоту $\nu^{(i)}$ гармонического компонента в соответствии с максимумом $R^{(i-1)}(\nu)$. Для устойчивости процесса вычитания спектра целесообразно в качестве оценки амплитуды колебания принять $c^{(i)} = g a[R^{(i-1)}, \nu^{(i)}]$, где g - коэффициент, обычно не превосходящий 1.

2) Образуем i-ый остаточный спектр

$$R^{(i)}(\nu) = R^{(i-1)}(\nu) - [c^{(i)} W_N(\nu - \nu^{(i)}) + (c^{(i)})^* W_N(\nu + \nu^{(i)})]. \qquad (3.4.11)$$

3) Рассмотрим $R^{(i)}$ и накопленные амплитуды $\{c^{(i)}\}$ с точки зрения проверки условия остановки. Если последнее не выполняется, продолжим итерации с шага 1; в противном случае перейдем к шагу 4. В качестве условия остановки могут, в частности, использоваться: а) малость пиков $R^{(i)}$ по сравнению с некоторым заданным наперед значением; б) достижение принятого числа итераций; в) существенное замедление скорости роста модулей амплитуд накопленных компонентов.

4) Поскольку найденный спектр $\{\nu^{(i)}, c^{(i)}\}$ искусственно образован так, что он имеет нулевое разрешение, выберем не-

которую сглаживающую функцию $B(\nu)$ и согласуем ее ширину с реальным временным окном. Например, можно принять для $B(\nu)$ гауссовский профиль с шириной $\sim T^{-1}$.

5) Образуем итоговый спектр $S(\nu)$ сворачиванием K компонентов $\{\nu^{(i)}, c^{(i)}\}$ с $B(\nu)$ и добавлением остаточного спектра:

$$S(\nu)=\sum_{i=1}^{K}\left[c^{(i)}B(\nu-\nu^{(i)})+(c^{(i)})^{*}B(\nu+\nu^{(i)})\right]+R^{(K)}(\nu). \qquad (3.4.12)$$

Дальнейшие подробкости, связанные с применением алгоритма CLEAN, а также примеры его использования при анализе искаженных скважностью данных можно найти в упомянутой работе Робертса и др. (1987). Типичное число требуемых итераций при $g \simeq 0,5$ составляет ~100. Небольшие неточности в определении частот максимумов нивелируются при последовательных приближениях. Как указывают Робертс и др. (1987), алгоритм CLEAN успешно работал в большинстве случаев, для которых проводилось его испытание.

Характеризуя алгоритм CLEAN в целом, необходимо отметить, что он эквивалентен описанному в предыдущем параграфе гармоническому отбеливанию. Как и последний, иногда CLEAN может оказаться неэффективным (например, если форма искомого колебания далека от синусоидальной). Вместе с тем применение CLEAN требует более мощных вычислительных средств и связано с некоторыми искусственными операциями, в частности, сглаживанием при помощи произвольной функции $B(\nu)$. По этим причинам, как нам представляется, следует отдать предпочтение давно применяющемуся методу отбеливания во временной области.

§ 3.5. Выделение тренда

Вопросы, связанные с исследованием систематического изменения временных рядов - *тренда* - намеренно были оставлены нами до тех пор, пока не будут изложены необходимые спектральные и статистические методы. Фактически речь идет об анализе низкочастотной области спектра мощности шириной порядка $10T^{-1}$, где T - общая продолжительность наблюдений. Даже в том случае, когда тренд сам по себе не представляет

интереса, его выделение необходимо для изучения спектра на высоких частотах; особенно тщательное рассмотрение должно быть предпринято для неравномерных рядов. Конечно, в общем случае не существует полностью объективных методов выделения тренда, в том числе и для равномерных рядов, однако варьирование процедуры с целью обеспечить устойчивость выводов обычно приводит к вполне разумным результатам.

Я.Пельт (частное сообщение) справедливо полагает, что нужно учитывать тренд на всех этапах анализа, так что общая его схема такова: 1) проводится предварительное выделение тренда; 2) анализируется ряд остатков; 3) моделируется суммарный ряд, включающий тренд и остатки; 4) анализируется ряд остатков от модели пункта 3 и т.д. Таким образом, моделировать следует *оригинальные* данные, учитывая всю имеющуюся информацию.

Линейный тренд. Прежде чем изучать проблему выделения тренда, относящуюся к классу *обратных* задач, полезно рассмотреть *прямую* задачу о влиянии простейших видов регулярного смещения на периодограмму.

Обратимся сначала к линейной модели тренда, заданной соотношением

$$f(t) = \begin{cases} \overline{f} + \dfrac{\Delta f}{T}\left(t - \dfrac{T}{2}\right), & 0 \leqslant t \leqslant T, \\ 0, & t < 0,\ t > T, \end{cases} \qquad (3.5.1)$$

где

$$\overline{f} = \frac{f(0)+f(T)}{2}, \quad \Delta f = f(T)-f(0) \qquad (3.5.2)$$

- соответственно среднее значение и приращение функции на интервале $[0, T]$. Отвлекаясь от несущественной пока дискретности отсчетов, найдем периодограмму $P_T(\nu)$ функции (1) согласно (1.6.6):

$$P_T(\nu) = \frac{1}{T^2}\left|\int_0^T e^{-i\cdot 2\pi\nu t}\, f(t)\,dt\right|^2 =$$

$$= \left(\overline{f}\,\frac{\sin\vartheta}{\vartheta}\right)^2 + \left[\frac{\Delta f}{2}\,\frac{1}{\vartheta}\left(\cos\vartheta - \frac{\sin\vartheta}{\vartheta}\right)\right]^2, \qquad (3.5.3)$$

где $\vartheta = \pi\nu T$. Вводя для удобства функцию

$$\mathrm{tr}(x) = \frac{1}{\pi}\frac{d}{dx}\,\mathrm{sinc}(x) = \frac{1}{\pi x}\left(\cos\pi x - \frac{\sin\pi x}{\pi x}\right), \qquad (3.5.4)$$

(см. Приложение 3), перепишем эту формулу в окончательном виде:

$$P_T(\nu) = \left[\overline{f}\ \mathrm{sinc}(\nu T)\right]^2 + \left[\frac{\Delta f}{2}\mathrm{tr}(\nu T)\right]^2 . \qquad (3.5.5)$$

Примечательно, что вклады в периодограмму, обусловленные смещением и наличием тренда, разделяются. При $\Delta f \to 0$ тренд исчезает, и мы приходим к знакомому из примера 1.3.2 результату:

$$P_T^{\mathrm{const}}(\nu) = [\overline{f}\ \mathrm{sinc}(\nu T)]^2. \qquad (3.5.6)$$

Периодограмма центрированной зависимости определяется вторым слагаемым в формуле (5). Функции $\mathrm{sinc}^2(x)$ и $\mathrm{tr}^2(x)$, характеризующие действие смещения и тренда, представлены на рис. 3.5.1. Относительная роль указанных факторов может быть задана параметром

$$\varkappa \equiv \left(\frac{\Delta f}{2\overline{f}}\right)^2 . \qquad (3.5.7)$$

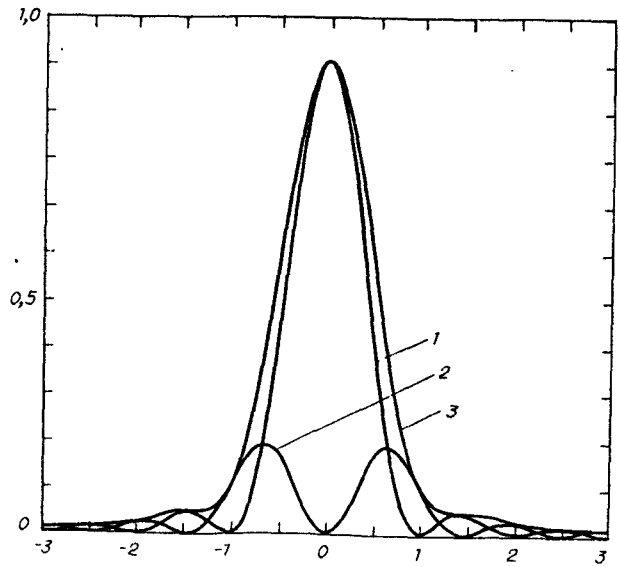


Рис. 3.5.1. Функции $\mathrm{sinc}^2(x)$, $\mathrm{tr}^2(x)$ и их сумма (указаны соответственно номерами 1, 2 и 3)

На рис. 3.5.1 приведена также периодограмма (5) при $\bar{f} = 1$, $\varkappa = 1$.

Весьма существенно, что основные максимумы $\mathrm{tr}^2(\nu T)$ расположены близко к центральному максимуму $\mathrm{sinc}^2(\nu T)$. Мы имеем для первой из этих функций

$$\nu_{max} \simeq \pm \frac{0,6626}{T}, \qquad (3.5.8)$$

в то время как ширина максимума второй функции составляет величину $\sim T^{-1}$. Как было описано в § 1.6, при дискретизации отсчетов спектр скважности «цепляется» за максимумы истинного спектра, так что в периодограмме центрированного ряда присутствуют все ложные максимумы на частотах, кратных частотам модуляции моментов наблюдения (например, кратных 1 циклу в сутки). С целью избавиться от этих максимумов ряды при поисках периодичности обычно центрируют и полагают, что для центрированных рядов прямое проникновение частот модуляции в периодограмму невозможно. Однако мы убедились выше, что и после вычитания из ряда среднего значения его периодограмма имеет максимумы на частотах (8), обусловленные трендом. Эти максимумы могут быть достаточно интенсивны, и тогда в периодограмме ложные максимумы на частотах, близких к частотам модуляции, не исчезают. Аналогичное заключение справедливо и для более сложных моделей тренда.

Таким образом, если ставится задача выделения скрытой периодичности, то после рассмотрения периодограммы исходного ряда $\{t_k, y_k\}$ следует рассчитать периодограмму ряда $\{t_k, y_k^{(o)}\}$, из которой при помощи метода наименьших квадратов устранен по крайней мере линейный тренд:

$$y_k^{(o)} = y_k - (\alpha_0 + \alpha_1 t_k), \quad k = 0, 1, \ldots, N-1,$$

$$\alpha_0 = \frac{\sum t^2 \sum y - \sum t \sum ty}{N \sum t^2 - (\sum t)^2}, \qquad (3.5.9)$$

$$\alpha_1 = \frac{N \sum ty - \sum t \sum y}{N \sum t^2 - (\sum t)^2}.$$

Заметим, что здесь не предполагается равномерность моментов наблюдения, а ряд $\{y_k^{(o)}\}$ автоматически получается центриро-

ванным.

Полиномиальный тренд. При значительной длине ряда линейное описание его систематического хода обычно становится недостаточным, и тогда наиболее употребительным является устранение тренда при помощи полинома

$$f(t) = \alpha_0 + \alpha_1 t + \ldots + \alpha_m t^m. \qquad (3.5.10)$$

Если для нахождения коэффициентов $\alpha_0, \ldots, \alpha_m$ используется метод наименьших квадратов, то нужно решить систему $2m+1$ нормальных уравнений

$$\begin{cases} \alpha_0 \sum t^j + \alpha_1 \sum t^{j+1} + \alpha_2 \sum t^{j+2} + \ldots + \alpha_m \sum t^{j+m} = \sum t^j y, \\ j = 0,\ 1,\ 2, \ldots,\ m\ , \end{cases} \qquad (3.5.11)$$

где суммирование производится по $k = 0, 1, \ldots, N-1$.

На практике устранение тренда целесообразно проводить, последовательно увеличивая порядок полинома m. По сути дела устранение тренда можно рассматривать как частный случай задачи отбеливания, когда отбеливанию подвергается область спектра, расположенная ниже интересующего нас диапазона частот. Ввиду этого при выборе оптимального порядка m_0 можно руководствоваться теми же соображениями, которые были изложены в предыдущем параграфе. Следует, в частности, тщательно следить за устойчивостью результатов относительно выбора m.

Сходство задач гармонического отбеливания и устранения тренда естественно приводит к представлению тренда тригонометрическим полиномом

$$f(t) = \alpha_0 + \sum_{n=1}^{m} \left[\alpha_n \cos(2\pi n \nu_0 t) + \beta_n \sin(2\pi n \nu_0 t) \right], \qquad (3.5.12)$$

где основная частота $\nu_0 \leqslant T^{-1}$. Вообще говоря, как коэффициенты $\{\alpha_0, \alpha_n, \beta_n\}$, так и частоту ν_0 можно выбрать оптимальным с точки зрения метода наименьших квадратов образом. В столь общей постановке задача сводится к решению сложной системы уравнений (см., в частности, Серебренников и Первозванский, 1965). Более простой подход опирается на предварительное

определение частоты ν_0 по виду периодограммы и дальнейшее нахождение коэффициентов полинома (12) из системы линейных уравнений, аналогичных (3.3.9). Опыт показывает, что при значительной скважности наблюдений, характерной для астрономических временных рядов, система нормальных уравнений нередко бывает плохо обусловлена, и тогда эффективны ограничения на амплитуду коэффициентов Фурье (Люси, 1976; Пельт, 1980). В указанных работах, а также статьях Джентлмена (1973, 1974), Флетчера и Пауэлла (1974) можно найти дальнейшие сведения, касающиеся численного представления в форме (12).

Скользящее среднее. Излагаемый в данном пункте метод выделения тренда применим только к равномерным рядам. Если из общего рассмотрения ряда ясно, что его регулярное поведение достаточно сложно и описание тренда ряда как целого потребует введения полинома высокого порядка, можно попытаться аппроксимировать ряд при помощи метода наименьших квадратов полиномами *по частям*: сначала первую группу $2l+1$ точек, затем группу точек той же длины, начиная со второй, и т.д. При этом в качестве тренда принимаются значения полиномов в средних точках групп.

Нахождение скользящего среднего подробно рассматривается во многих руководствах (см. Кендалл, 1976; Андерсон, 1971), поэтому мы приведем лишь основной результат и частную формулу. Именно, можно показать, что описанная выше процедура нахождения тренда сводится к вычислению средневзвешенных значений

$$f(t_k) = \sum_{j=-l}^{l} \gamma_j^{(m)} y_{k+j} \qquad (3.5.13)$$

по $2l+1$ отсчетам ряда с весами $\gamma_j^{(m)}$, зависящими от порядка m аппроксимирующего полинома. Например, при усреднении групп по семь точек с помощью полиномов 3-го порядка коэффициенты $\gamma_j^{(3)}$, $|j| \leqslant 3$ равны соответственно

$$\frac{1}{21}\,[-2,\ 3,\ 6,\ 7,\ 6,\ 3,\ -2]\ , \qquad (3.5.14)$$

так что

$$f(t_k) = \frac{1}{21}\,(-2y_{k-3}+3y_{k-2}+6y_{k-1}+7y_k+$$

$$+6y_{k+1}+3y_{k+2}-2y_{k+3}),\quad k = 3,\ldots,\ N-4. \qquad (3.5.15)$$

Сумма весов $\gamma_j^{(m)}$ всегда равна 1, они симметричны, $\gamma_{-j}^{(m)} = \gamma_j^{(m)}$, и для полиномов четного порядка $m = 2r$ коэффициенты будут те же, что и для полиномов порядка $2r+1$. Сводка весов $\gamma_j^{(m)}$, $|j| \leqslant l$ для $m = 3,5$ и $2l+1 = 5,7,\ldots,19,21$ имеется в монографии Кендалла (1976).

Обобщение для неравномерных рядов. Обобщение описанной выше процедуры легко реализуется численным образом (В. Теребиж, А. Теребиж и Бирюков, 1989). Именно, вместо аппроксимации полиномами заданного количества отсчетов можно аппроксимировать полиномами все отсчеты, попадающие в заданные интервалы времени, последовательно смещенные друг относительно друга.

Пусть, например, все моменты наблюдения сосредоточены в промежутке $[0,T]$, а сглаживание производится по интервалам длительности τ, последовательно смещаемым на $\delta t \leqslant \tau$. В каждом интервале сглаживания через наблюдаемые отсчеты методом наименьших квадратов проводится полином заданной степени m, отвечающей характеру нужного усреднения, и затем в качестве отсчета сглаженного ряда принимается значение полинома в средней точке. Если крайние интервалы сглаживания выбраны так, что их центры совпадают с моментами 0 и T, то в итоге мы получим $N = [T/\delta t]+1$ значений равномерного временного ряда с шагом δt.

Вообще говоря, можно принять δt сколь угодно малой величиной и получить сглаженный ряд, сколь угодно близкий к непрерывному процессу, однако отличному от исходного непрерывного процесса. Действительно, если исходный временной ряд был близок к равномерному, т.е. *коэффициент вариации* интервалов между измерениями $C(\Delta t) \equiv \sigma(\Delta t)/\overline{\Delta t}$ мал, то по теореме Котельникова-Шеннона ряд не содержит информации о реальном процессе в области частот, превосходящих $\nu_{max} \approx (2\overline{\Delta t})^{-1}$. Поскольку высокочастотная часть спектра недоступна изна-

чально, нет смысла выбирать значения δt, намного меньшие $\overline{\Delta t}$. В противоположном случае существенно неравномерного исходного ряда, когда $C(\Delta t) \gtrsim 1$, минимальное полезное значение δt находится приблизительно в области наименьших интервалов времени между измерениями.

Вопрос о выборе степени сглаживающего полинома m и ширины временного окна τ не изучен надлежащим образом. Некоторые полезные соображения для случая равномерного ряда имеются в книге Кендалла и Стюарта (1966); следует, впрочем, заметить, что при обсуждении эффекта Слуцкого-Юла в гл. 46 упомянутые авторы не указывают, что периодичность при сглаживании случайного ряда возникает лишь при *одновременном* образовании разностей, т.е. введении как низкочастотной, так и высокочастотной фильтрации. Вычисление скользящего среднего (13) эквивалентно в частотной области введению фильтра с пропусканием $K_{m,\tau}(\nu)$, зависящим только от параметров m и τ. Аналогичная операция для неравномерного ряда приводит к низкочастотной фильтрации, зависящей и от самих отсчетов временного ряда. Тем не менее при должном подходе сглаживание неравномерного ряда «скользящим полиномом» чрезвычайно полезно как при выделении медленных колебаний процесса, так и при изучении области высоких частот путем вычитания сглаженного ряда из исходного.

Как легко видеть, последовательное проведение сглаживания скользящим полиномом, вычитания сглаженного ряда и вычисления спектра остаточного ряда составляет операцию, названную в § 3.3 отбеливанием. Но в данном случае отбеливаются не гармонический или полигармонический компоненты (дискретный спектр), а вся низкочастотная область спектра. Например, если максимум в области высоких частот значительно уменьшился после такого низкочастотного отбеливания, можно полагать, что этот максимум не присущ спектру изучаемого процесса, а обусловлен лишь распределением моментов наблюдения. При полиномиальном отбеливании следует руководствоваться соображениями, аналогичными изложенным в § 3.3.

Пример низкочастотного полиномиального отбеливания приведен в следующем параграфе.

Статистическая значимость тренда устанавливается на ос-

нове среднеквадратических отклонений вычисленных значений коэффициентов и оценок доверительных интервалов. В простых случаях можно воспользоваться известными теоретическими результатами, в более сложных - *методом Монте-Карло*, совпадающим в данном контексте с *принципом рандомизации Фишера* или так называемым *методом бутстрэпа* (bootstrap); превосходное популярное изложение последнего дано Диаконисом и Эфроном (1983).

Классический способ оценки реальности тренда для ряда $\{y_k\}$ основывается на вычислении *статистики Аббе* (см. Кроу и др., 1960; Большев и Смирнов, 1965; Бэрк и др., 1978). Пусть

$$\sigma^2_{tot} = \frac{1}{N-1} \sum_{k=0}^{N-1} (y_k - \bar{y})^2 \qquad (3.5.16)$$

- обычная несмещенная оценка дисперсии ряда, а

$$\sigma^2_{in} = \frac{1}{2(N-1)} \sum_{k=0}^{N-2} (y_{k+1} - y_k)^2 \qquad (3.5.17)$$

- оценка его внутренней дисперсии, сохраняющейся и при устранении медленного систематического изменения ряда. Очевидно, при наличии тренда отношение

$$A_N = \frac{\sigma^2_{in}}{\sigma^2_{tot}} = \frac{1}{2} \frac{\sum_{k=0}^{N-2} (y_{k+1} - y_k)^2}{\sum_{k=0}^{N-1} (y_k - \bar{y})^2} \qquad (3.5.18)$$

будет мало по сравнению с 1. С другой стороны, при совершенно случайных колебаниях членов ряда около среднего ожидается $A_N \sim 1$. Поэтому наблюдаемая на опыте величина A_N^{obs} служит показателем степени выраженности тренда. Функция распределения параметра A_N в случае нормальной совокупности $\{y_k\}$ была впервые изучена фон Нейманом (1941) и затем многими авторами (см. § 4.1). Критические значения $A_{N,q}$ отвечающие принятому уровню значимости q, даны в Приложении 6.

§ 3.6 Фликкер-шум. Переменность блеска ядер сейфертовских галактик

Характерная особенность квазаров и ядер сейфертовских галактик (свойства активных ядер галактик обсуждаются в

§ 2.8) заключается в том, что оптическая светимость этих объектов обнаруживает относительные колебания порядка единицы во всем изученном диапазоне временных масштабов. Иллюстрацией к сказанному может служить рис. 3.6.1, на котором представлены вариации потока в полосе *B* ядра сейфертовской галактики NGC 4151 за 20-летний период с 1968 г. по 1987 г. (общее количество фотоэлектрических измерений $N = 462$, подробнее см. ниже). С первого взгляда на рис. 3.6.1 создается впечатление, что здесь присутствуют три независимых компонента: общий тренд с характерным временем порядка 10 лет, колебания с длиной цикла порядка нескольких лет и быстрые вспышки протяженностью около 10 дней. В ранних работах такие компоненты считались реальными, а для второго из них предполагалась и квазипериодичность. При наличии данных более длительных наблюдений активных ядер (см., например, сводку Энжиона и Смита, 1985, для 3C 273) можно выделить и значительные колебания светимости с длиной цикла в десятки лет.

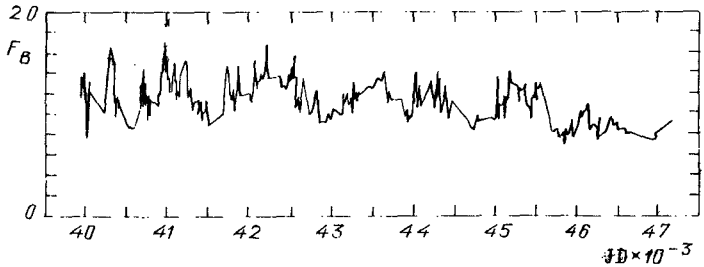


Рис. 3.6.1. Кривая блеска ядра галактики NGC 4151 в полосе *B* с 1968 по 1987 г. В качестве единичного выбран поток от звезды 15^m

Переменность в широком диапазоне временных масштабов присуща не только активным ядрам галактик, и даже не только многим объектам, изучаемым современной астрономией, но проявляется в самых разнообразных физических явлениях: шуме электронных ламп и гранулированных сопротивлений, колебаниях скорости подводных течений в океанах, флуктуациях громкости звука в радиопередачах, течении песка в песочных часах и т.п. В спектральной области такое поведение процессов отвечает степенному убыванию спектральной плотности в широ-

ком диапазоне частот:

$$g(\nu) = \text{const}\ \nu^{-\gamma}, \qquad (3.6.1)$$

где показатель γ лежит обычно в пределах от 0,6 до 2,0 (Малахов, 1968; Рытов, 1976). Случайный процесс со спектром вида (1) называется *фликкер-шумом*. Насколько нам известно, впервые аналогия между флуктуациями блеска активных ядер галактик и классическим фликкер-шумом рассматривалась в очень интересной работе Пресса (1978). В отношении перечисленных выше «компонентов» кривой блеска NGC 4151 указанная аналогия может означать, что в действительности каких-либо выделенных масштабов переменности нет, и наблюдаемые колебания отражают лишь свойства конкретной реализации *единого по своей природе* процесса выделения энергии центральным источником.

Весьма примечательно, что сложившаяся ранее картина переменности потока активных ядер в рентгеновском диапазоне спектра радикально изменилась с получением длительных (до 80 часов) непрерывных рядов измерений при помощи счетчиков, установленных на спутнике EXOSAT (Лоуренс и др., 1987; Мак Харди и Черны, 1987; Машоцки, 1987; Урри, 1988). Выяснилось, что для этих объектов свойственны не плавные вариации, изредка прерываемые вспышками, а непрерывные флуктуации значительной амплитуды со спектром мощности $g(\nu) \propto \nu^{-1}$ в области частот 10^{-5}- 10^{-3}Гц. Мак Харди и Черны (1987) указали, что рентгеновские кривые блеска NGC 4051 и NGC 5506 напоминают введенные Мандельбротом (1977) *фракталы* - объекты, структура которых сохраняется при изменении масштабов (известный пример фрактала - скалистая береговая линия).

В данном параграфе мы сначала кратко обсудим, отчасти следуя цитированным выше авторам, теоретические аспекты фликкер-шума, а затем рассмотрим с этой точки зрения переменность оптической светимости наиболее хорошо исследованной сейфертовской галактики NGC 4151 (см. также § 2.8 и статью В.Теребижа, А.Теребиж и В.Бирюкова, 1989).

Фликкер-шум. Сопоставим три вида случайных процессов: белый шум (WN), случайные блуждания (RW) и фликкер-шум

(FN). Типичные реализации этих процессов представлены на рис. 3.6.2. Выше уже отмечалось, что визуальный анализ способен обнаружить весьма глубокие свойства случайных процессов; как показывают численные эксперименты (Пресс, 1978), человек легко разделяет процессы, отличающиеся только при обращении к трех- или четырехточечным характеристикам. Различие внутренней структуры процессов на рис. 3.6.2 представляется слишком большим, чтобы его можно было назвать лишь видимым «невооруженным глазом».

Поскольку спектральная плотность белого шума (1.5.16) не убывает, по крайней мере теоретически с ростом частоты, ему присущи сколь угодно быстрые колебания. С другой стороны, в области низких частот его плотность $g_{WN}(\nu)$ ограничена, так что среднее значение белого шума при рассмотрении все более длинных интервалов времени устойчиво. Формально эти свойства связаны с расходимостью интегральной мощности

$$\sigma^2(\nu_1, \nu_2) = \int_{\nu_1}^{\nu_2} g(\nu)\,d\nu \qquad (3.6.2)$$

белого шума при $\nu_2 \to \infty$ и сходимостью при $\nu_1 \to 0$ (ср. с формулой (1.5.14)).

Белый шум $\xi_{WN}(t)$ принимает в каждый данный момент времени значение, совершенно не связанное с его поведением в предыдущие моменты. Предположим, что вдоль прямой линии случайно двигается частица, смещения которой описываются процессом белого шума. Тогда суммарное расстояние, пройденное частицей к моменту t, равно

$$\xi_{RW}(t) = \int_0^t \xi_{WN}(t')\,dt' \ . \qquad (3.6.3)$$

Функция $\xi_{RW}(t)$ задает *процесс случайного блуждания*, являющийся математической моделью *броуновского движения* мелких частиц в жидкости. Реализация на рис. 3.6.2, *б* была получена из реализации 3.6.2, *а* именно с помощью соотношения (3). Текущая спектральная плотность процесса случайных блужданий на интервале [0, t] легко находится из (1.5.43) и (1.5.44); она равна

$$g_{RW}(\nu, t) = \left[\frac{\sin(\pi\nu t)}{\pi\nu}\right]^2 g_{WN}. \qquad (3.6.4)$$

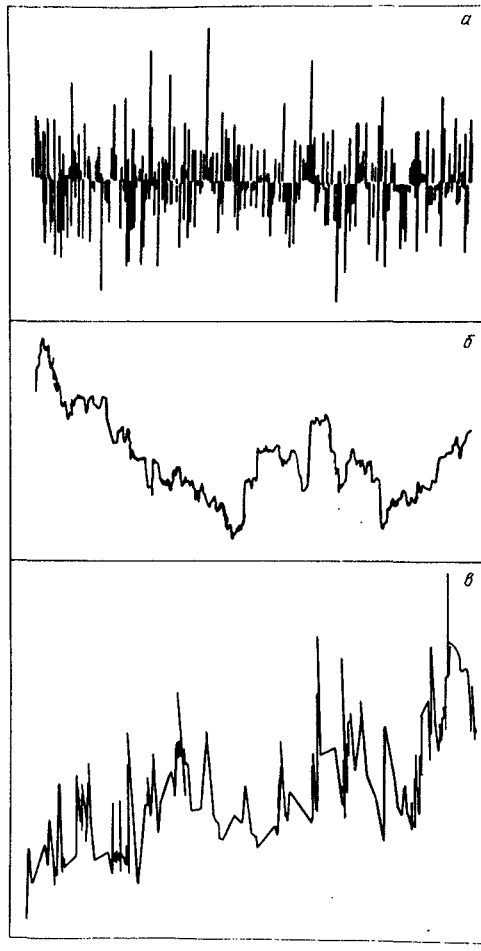


Рис. 3.6.2. Реализации белого шума (а), процесса случайного блуждания (б) и фликкер-шума (в)

Поскольку плотность зависит от времени, процесс нестационарен. Воспользовавшись соотношением (П.3.4), находим, что полная мощность

$$\sigma^2_{RW}(t) = \int_{-\infty}^{\infty} g_{RW}(\nu, t)\,d\nu = g_{WN}\ t\,. \qquad (3.6.5)$$

Таким образом, с ростом t полная мощность процесса случайных блужданий расходится. Это связано с расходимостью низкочастотной части спектральной плотности (4), так как для

высоких частот плотность убывает как ν^{-2}. Во временной области такое поведение спектральной плотности эквивалентно отсутствию быстрых хаотических колебаний $\xi_{RW}(t)$ значительной амплитуды и наличию длинных «волн» отклонений от среднего, описываемых *законами арксинуса* (Феллер, 1966, гл. 3). Текущее среднее значение процесса случайных блужданий ввиду «красной расходимости» плохо определено.

Наибольшее сходство с действительными изменениями блеска активных ядер галактик имеет реализация фликкер-шума на рис. 3.6.2, *в*. Эта реализация содержит и быстрые «вспышки», и плавные волны, как будто указывающие на периодичность с большой длиной цикла (в связи с долговременными колебаниями фликкер-шума Пресс заметил: «какой теоретик сможет устоять перед искушением интерпретировать их?»).

Для сходимости полной мощности $\sigma^2(\nu_1, \nu_2)$ фликкер-шума нужно, чтобы его спектральная плотность убывала при $\nu \to \infty$ быстрее, чем ν^{-1}, а при $\nu \to 0$ возрастала медленнее, чем ν^{-1}. Классический фликкер-шум с плотностью $g_{FN}(\nu) \propto \nu^{-1}$ характеризуется логарифмической расходимостью $\sigma^2(\nu_1, \nu_2)$ как при $\nu_2 \to \infty$, так и при $\nu_1 \to 0$. Это означает, что для него типичны и большие резкие флуктуации (как для белого шума), и плавные уходы большой амплитуды от среднего значения (как для случайных блужданий). Другими словами, фликкер-шум обнаруживает заметную переменность в широком диапазоне временных масштабов (при $\gamma = 1$ на каждую декаду частот приходится одна и та же мощность).

В этой связи представляет интерес замечание Пресса (1978), касающееся часто встречающихся в астрономической литературе утверждений о «квазипериодичности» кривых блеска разных объектов. Глаз человека обладает свойством выделять из шума колебания, даже если они не значимы статистически. Для шума с расходящейся на низких частотах спектральной плотностью наибольшая мощность сосредоточена в области периодов, сравнимых с длиной всего ряда. Но глаз эффективно устраняет и линейный тренд (вследствие привычки отстраиваться от перспективы?), так что гармоники с очень большими периодами исключаются. Поэтому представляется весьма прав-

доподобным, что наиболее заметными для глаза будут колебания с периодом порядка 1/3 длины ряда.

В связи с широкой распространенностью фликкер-шума его происхождение неоднократно обсуждалось как в физической, так и в технической литературе, и тем не менее приходится признать, что механизм генерации фликкер-шума почти во всех случаях остается неясным.

Известно (см., например, Малахов, 1968), что степенной спектр мощности получается как частный случай в рамках модели дробового шума (2.8.1), а именно:

$$\xi(t) = \sum_j a_j s\left(\frac{t-t_j}{\theta_j} \right), \qquad (3.6.6)$$

где $a_j > 0$ и θ_j - соответственно случайные амплитуды и длительности вспышек детерминированной формы $s(u)$, а $\{t_j\}$ - пуассоновская последовательность моментов вспышек. Если считать протяженность вспышек фиксированной величиной ($\theta_j \equiv \theta = \text{const}$), то спектральная плотность $\xi(t)$ согласно (2.8.5) равна

$$g_\xi(\nu) = \bar{n}\,\overline{a^2}\,\theta^2\,|S(\nu\theta)|^2, \quad \nu > 0, \qquad (3.6.7)$$

где $\bar{n}$ - средняя частота вспышек, $S(z)$ - образ Фурье кривой блеска отдельной вспышки. Предположим, что «хвост» вспышки имеет степенную форму, т.е. $s(u) \propto |u|^{-q}$, $0 < q < 1$ при $|u| \gg 1$. Тогда $g_\xi(\nu)$ имеет при $\nu \ll \theta^{-1}$ вид (1) с показателем

$$\gamma = 2(1-q), \qquad 0 < \gamma < 2. \qquad (3.6.8)$$

Для того чтобы получить $\gamma \simeq 1$, нужно принять $q \simeq 1/2$, т.е. вспышки затухают весьма медленно, как $t^{-1/2}$ при $t \gg \theta$. Именно наложение протяженных «хвостов» вспышек обеспечивает в данной модели сильную низкочастотную переменность, отвечающую фликкер-шуму. Но вместе с тем, как показывает (2.8.5), это наложение в случае $q \leqslant 1$ приводит и к расходимости среднего значения $\langle\xi(t)\rangle$, т.е. процесс (6) становится нестационарным (дисперсия σ_ξ^2 конечна при $q > 0{,}5$). Сама по себе расходимость среднего не является трудностью,

так как вариации светимости активного ядра не обязаны быть стационарными, однако при этом может возникнуть проблема согласования среднего значения и дисперсии переменной составляющей.

Соотношение (8) показывает, что при $q \to 0$ ($\gamma \to 2$) мы приходим к чисто диффузионному закону (5), когда дисперсия линейно растет со временем, а при $q \to 1$ ($\gamma \to 0$) - к белому шуму.

Вообще говоря, фликкер-шум можно образовать и задавая последовательность вспышек произвольной формы, если рандомизировать их длительности θ_j. Тогда в соответствии с последним из соотношений (2.8.5) спектральная плотность

$$g_\xi(\nu) = \bar{n}\,\overline{a^2} \int_0^\infty |S(\nu\tau)|^2\, p_\theta(\tau)\tau^2 d\tau\,, \qquad (3.6.9)$$

где $p_\theta(\tau)$ - плотность распределения θ_j. Полагая $p_\theta(\tau) \propto \tau^{-2}$, $\tau_0 \leqslant \tau \leqslant \tau_1$, находим из (9), что $g_\xi(\nu) \propto \nu^{-1}$ в области частот $(2\pi\tau_1)^{-1} \leqslant \nu \leqslant (2\pi\tau_0)^{-1}$. На более низких частотах $g_\xi(\nu)$ соответствует белому шуму, а на высоких - спадает пропорционально ν^{-2}. Конечно, предположение о том, что параметр затухания θ_j подчиняется специальной плотности распределения в широком диапазоне значений, представляется довольно искусственным (Пресс, 1978).

Анализ спектральной плотности. Основой для анализа переменности блеска ядра галактики NGC 4151 послужили фотоэлектрические измерения с диафрагмой диаметром 25 - 27" в полосах *U* и *B*, выполненные Пенстоном и др. (1971, 1974), Белоконь (1989) и Лютым (1989). Количество отсчетов в полосе *U* составило 388, в *B* - 462. Измерения с указанными двумя фильтрами анализировались единообразно; в *U* более высокий контраст ядра позволяет четче проявиться его вспышечной активности, в то время как оценки в *B* многочисленнее. Раздельное рассмотрение данных во всех отношениях привело к сходным между собой выводам, поэтому здесь мы представим результаты, полученные для синей области спектра.

Оценки звездных величин были трансформированы в потоки $F = \mathrm{dex}[-0,4(m-15)]$. Относительные вариации потоков $\sigma_F/\bar{F}$ составили ~30 % и 20 % для *U* и *B* соответственно Средний интервал времени между измерениями в *B* - 15,7 суток.

Спектр скважности ряда $\{F_B\}$ в области $\nu \leqslant 0,01$ цикл/сут представлен на рис. 3.6.3, *а*. Максимум на частоте 0,00273 цикл/сут отвечает годичной модуляции моментов наблюдения.

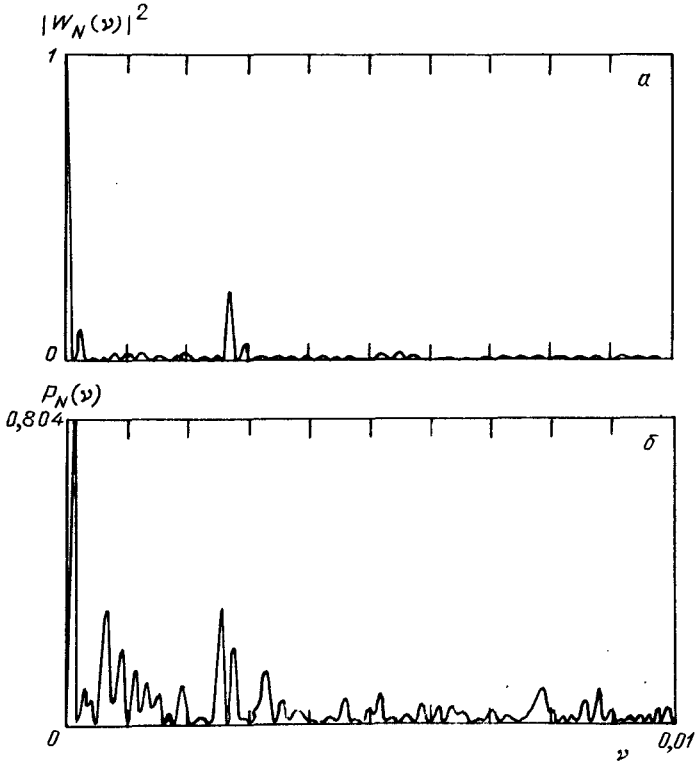


Рис. 3.6.3. Спектр скважности (а) и периодограмма Шустера (б) ряда, представленного на рис. 3.6.1

В области высоких частот имеется аналогичный максимум, соответствующий модуляции с периодом лунного месяца.

На рис. 3.6.3, *б* приведена периодограмма центрированного ряда $\{F_B\}$ для периодов $P \geqslant 100$ сут. Наибольший максимум расположен на частоте 10^{-4} цикл/сут ($P \simeq 27$ лет), он соответствует описанному выше тренду. Не исключено, что группа максимумов в области частот $5 \cdot 10^{-4}$ - $2 \cdot 10^{-3}$ цикл/сут (5,5 - 1,5 года) отвечает участку спектра с монотонно убывающей от низких частот плотностью, а провал при $\nu \simeq (2 \div 6) \cdot 10^{-4}$ цикл/сут может быть случайным эффектом реализации.

«Изрезанность» оценки спектра мощности не должна смущать, поскольку она является неотъемлемым свойством периодограммы Шустера (§ 2.3). Два пика $P_N(\nu)$ при $\nu \approx 2{,}7 \cdot 10^{-3}$ цикл/сут ($P \simeq 370$ сут) с очевидностью обусловлены трендом и годичной модуляцией моментов наблюдения, проявляющейся на спектре скважности. Действительно, эти пики исчезают, если построить периодограмму детрендированного ряда. На более высоких частотах вплоть до 0,1 сут (эта часть периодограммы не приводится) отсутствуют сколько-нибудь заметные особенности. Для контроля мы рассчитали LS-спектр колебаний блеска NGC 4151 (§ 3.1). Результирующая оценка спектра мощности оказалась практически совпадающей с той, которую дает периодограмма Шустера.

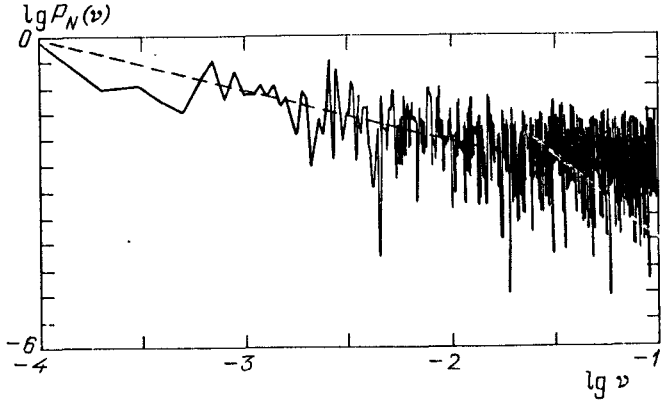


Рис. 3.6.4. Периодограмма кривой блеска NGC 4151 в двойном логарифмическом масштабе. Штриховая прямая соответствует показателю $\gamma = 1$

Общий характер зависимости спектра мощности от частоты наглядно выявляется при построении периодограммы в двойном логарифмическом масштабе (рис. 3.6.4). Видно, что при $\lg \nu \leqslant -1{,}7$ ($P \geqslant 50$ сут) спектр мощности удовлетворительно описывается степенным законом (1) с показателем $\gamma \approx 1{,}0$. Это означает, что низкочастотные вариации светимости ядра галактики близки к фликкер-шуму. В качественном виде это заключение в отношении активных ядер галактик было сделано Прессом (1978) и в отношении NGC 4151 - Пахольчиком и др. (1983). Для частот выше $\lg \nu \approx -1{,}7$ спектр NGC 4151 плоский;

здесь, впрочем, имеющиеся данные становятся недостаточными для получения надежных выводов.

Итак, имеются две возможности: а) спектральная плотность вариаций блеска NGC 4151 плавно уменьшается с ростом частоты и нет свидетельств в пользу реальности отдельных компонентов; б) провал в области частот $(2\div6)\cdot10^{-4}$ цикл/сут не случаен, так что реальны по крайней мере два таких компонента.

Для дальнейшего изучения низкочастотной области спектра кривые блеска NGC 4151 в полосах U и B были сглажены в пределах интервала длиной 1000 сут, перемещавшегося с шагом 100 сут (см. § 3.5). Отсчеты в пределах каждого положения интервала аппроксимировались полиномом 4-го порядка, а затем в качестве сглаженной оценки потока принималось значение полинома в средней точке. Полученный таким образом равномерный ряд, включающий 71 точку, анализировался при помощи метода максимума энтропии так же, как это было сделано в § 2.8 для 3C 273. Была проверена и устойчивость результатов относительно изменения параметров процедуры сглаживания.

Мы укажем здесь лишь основные результаты проведенного анализа. Область спектра с частотами $\nu \leqslant 0{,}002$ цикл/сут мало изменилась в результате сглаживания, в то время как интенсивность высокочастотного участка спектра сильно уменьшилась. Подбор авторегрессионных моделей показал широкий минимум окончательной ошибки предсказания Акаике FPE(p) при небольших значениях порядка модели $p \simeq 2\div4$. Соответствующие этим порядкам кривые спектральной плотности свидетельствуют в пользу единого механизма колебаний в области $P \geqslant 500$ сут, т.е. не дают оснований выделить тренд в качестве особого компонента.

Более сильный аргумент в пользу единой природы низкочастотных и высокочастотных колебаний блеска следует из совместного рассмотрения данных в U и B. На рис. 3.6.5 представлены соотношения F_U- F_B отдельно для медленного компонента вместе с постоянной составляющей от ядра галактики «s + g» и быстрого компонента «f». Первый был выделен при помощи описанного выше сглаживания, второй - как раз-

ность между наблюдаемыми отсчетами и компонентом «$s + g$».

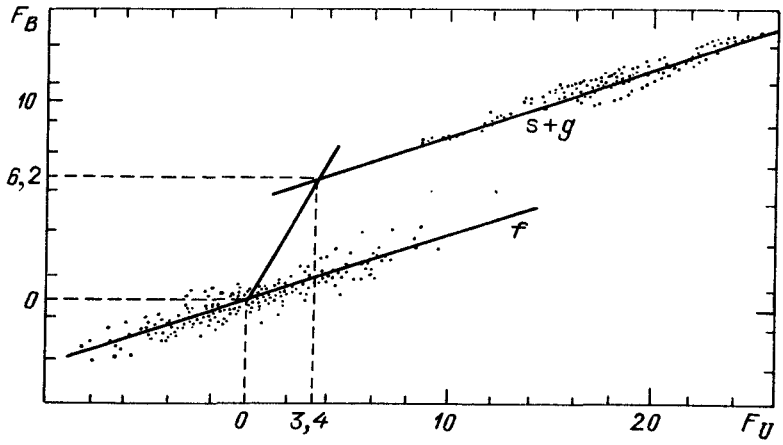


Рис. 3.6.5. Диаграмма F_U – F_B для медленных ($s+g$) и быстрых (f) вариаций блеска NGC 4151

Прежде всего обратим внимание, что наклоны зависимостей для обоих компонентов одинаковы в пределах точности измерений. Значения углов наклона согласуются с предположением о равенстве показателей цвета медленного и быстрого переменных компонентов: $(U - B)_{var} = -1,11 \pm 0,01$. К аналогичному выводу пришел ранее Гаген-Торн (1985, 1987), разделивший «на глаз» кривую блеска галактики на зоны повышенной и пониженной светимости. Рис. 3.6.5 демонстрирует равенство показателей цвета значительно более четко. Оно свидетельствует о том, что разделение на быстрый и медленный компоненты, вероятно, носит искусственный характер.

Если отказаться от указанного разделения, то на диаграмме F_U – F_B останется, очевидно, лишь несколько более размытая совокупность точек «$s + g$». Принимая для центральной области галактики $(U - B)_g = 0,66$, можно определить вклад постоянной звездной составляющей в суммарные потоки в пределах 27"-диафрагмы. Такое разделение дает (см. рис. 3.6.5): $F_U^{(g)} \approx 3,4$, $F_B^{(g)} \approx 6,2$. Таким образом, из всего среднего значения $\overline{F}_B = 11,4$ на долю переменного центрального источника приходится приблизительно половина потока: $\overline{F}_B^{var} \approx 5,2$.

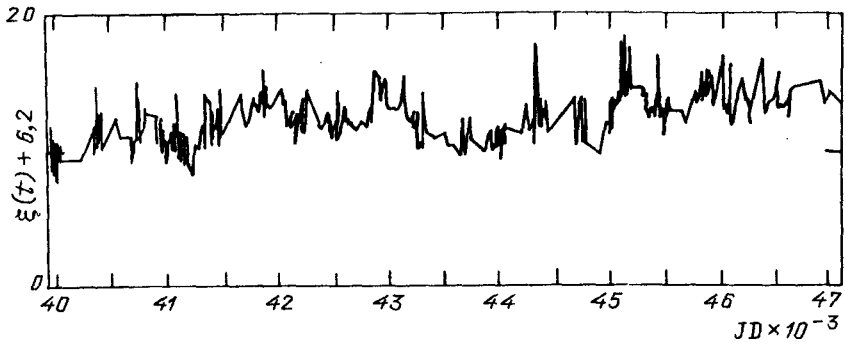


Рис. 3.6.6. Одна из реализаций случайного процесса, заданного соотношением (6)

Наконец, при помощи метода Монте-Карло были получены реализации фликкер-шума (6) как при фиксированной, так и при случайной продолжительности вспышек. В первом случае мы приняли $s(u) = (1 + u)^{-0,55}$ для $u \geqslant 0$. Во втором - $\bar{n} = 1$ сут$^{-1}$, $\bar{a} = 2/3$, $\tau_0 = 1$ сут, $\tau_1 = 3 \cdot 10^4$ сут, форма вспышки и плотность распределения амплитуд имели экспоненциальный вид (при этом средняя длительность вспышки $\bar{\theta} \approx 10$ сут). Моменты наблюдения процесса $\xi(t)$ были выбраны совпадающими с наблюдаемой совокупностью таковых для NGC 4151, значения ординат сдвинуты в положительном направлении на 6,2 в соответствии с найденной оценкой вклада постоянной звездной составляющей галактики. Реализованные кривые блеска при фиксированной продолжительности вспышек показали близкое сходство с наблюдаемой кривой блеска NGC 4151, однако при этом едва ли возможно согласовать дисперсию потоков $\sigma_F^2 \approx 4$ и низкое среднее значение переменной составляющей $\bar{F}_B^{\,var} \approx 5$. Типичная реализация для случайных θ_j представлена на рис. 3.6.6, ее периодограмма в обычном и двойном логарифмическом масштабах показана на рис. 3.6.7.

Обсуждение результатов. Модельная кривая блеска на рис. 3.6.6 во многих чертах напоминает реальную кривую блеска NGC 4151. Принимая во внимание упомянутое свойство глаза выделять не только двухточечные характеристики процесса, но и характеристики высоких порядков, сходство кривых блеска не следует недооценивать.

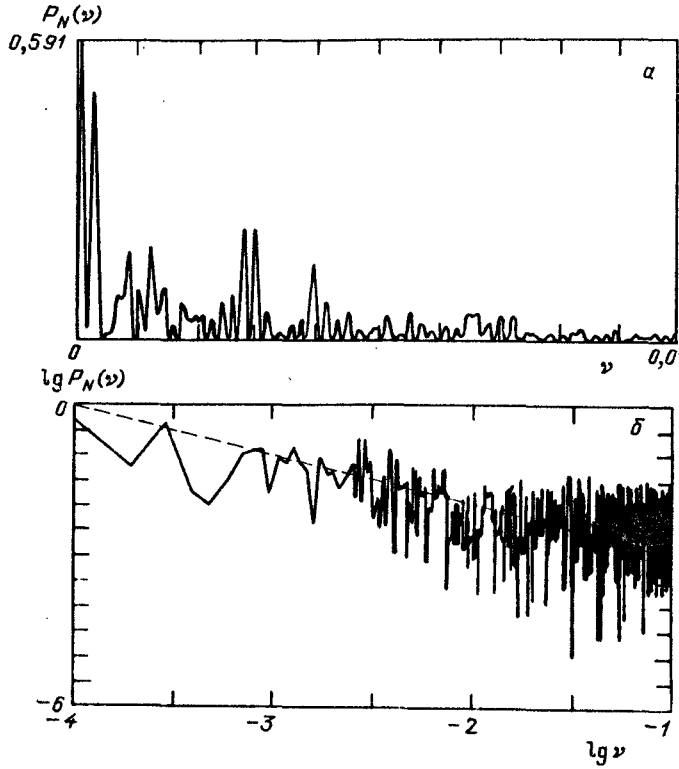


Рис. 3.6.7. Периодограмма модельной кривой блеска, приведенной на рис. 3.6.6, в линейном (а) и двойном логарифмическом (б) масштабах. Штриховая прямая соответствует наклону −1

Периодограмме случайной реализации (рис. 3.6.7) присущи столь же нерегулярные флуктуации, которые наблюдаются на периодограмме действительной кривой блеска (рис. 3.6.3). В первом случае этим флуктуациям заведомо не отвечают никакие периодические компоненты; то же справедливо, по-видимому, и для действительных колебаний блеска NGC 4151 с характерным временем, превосходящим десятки суток. Этот вывод согласуется и с указанными выше результатами регрессионного анализа. Для изучения более быстрых колебаний нужно выполнить специальную программу наблюдений.

В модели дробового шума (6) важную роль играет безразмерный параметр $N_0 \equiv n\bar{\theta}$ - среднее количество вспышек, слу-

чившихся в течение промежутка времени, равного средней продолжительности одной вспышки. Для согласования модели с данными наблюдений нам потребовалось принять $N_0 \sim 10$, т.е. допустить значительное перекрытие отдельных вспышек. Это согласуется с тем фактом, что при детальном изучении кривой блеска NGC 4151 в большом масштабе не удается выявить сколь-нибудь правильных отдельных вспышек. В частности, в исследованном частотном диапазоне возможная временная асимметрия процесса (*стрела времени*) не проявляется.

Успех модели со случайными θ_j во многом определяется введением степенного закона $p_\theta(\tau) \propto \tau^{-2}$; конкретная форма отдельных вспышек и распределение их амплитуд не имеют значения. Поэтому можно сказать, что трудность интерпретации перенесена в этой модели в другую область - необходимо понять, почему степенной закон распределения θ_j сохраняется в пределах широкого временного диапазона, охватывающего 3-4 порядка. Как отмечалось, сделанное предположение не представляется естественным, однако нельзя и полностью исключить возможность того, что близкий к степенному закон распределения продолжительностей вспышек реализуется в действительности. С другой стороны, столь же удовлетворительного согласия с имеющимися данными можно достичь и в модели с фиксированным θ, предполагая лишь медленное убывание «хвостов» вспышек по степенному закону. Сейчас трудно отдать решительное предпочтение одной из этих моделей.

Во введении к данному параграфу мы упомянули, что кривые блеска активных ядер галактик в различных диапазонах спектра обнаруживают свойство *масштабной инвариантности*, относящее их к классу одномерных фракталов. Автоподобие фрактальной кривой приводит к степенному возрастанию ее длины L по мере уменьшения длины интервала Δt, с которым проводятся измерения: $L \propto \Delta t^{-D}$. Рэйн (см. работу Мак Харди и Черны, 1987) показал, что *фрактальная размерность* D связана с показателем γ степенного спектра (1) простым соотношением

$$\gamma = 2(1 - D). \qquad (3.6.10)$$

В отличие от рентгеновских измерений, оценка D по оптическим данным затруднена ввиду неравномерности отсчетов. Пред-

варительное изучение показывает, что для оптических вариаций потока NGC 4151 степенной характер зависимости L от Δt выполняется. Здесь необходимо детальное исследование. Пока обратим лишь внимание на то, что в модели суперпозиции вспышек степенной формы мы имеем на основании (8) и (10) простое соотношение $D = q$, т.е. фрактальная размерность совпадает с показателем скорости угасания вспышек.

Суммируя результаты анализа, можно сказать, что кривая блеска NGC 4151 не содержит указаний на реальность отдельных компонентов с характерными временами, превосходящими 50 суток. Весьма правдоподобно, что наблюдаемые вариации потока порождаются наложением отдельных независимых вспышек.

§ 3.7. Аналогии в оптике

Уже Шустер (1906) отметил глубокое сходство между анализом временных рядов и спектральным анализом излучения. В основе обоих подходов лежат, по существу, идентичные понятия: распределение энергии излучения по частотам представляет собой спектральную плотность электромагнитных колебаний, рассматриваемых как стационарный случайный процесс (см. § 5.5). Ввиду большого опыта, накопленного спектроскопией, естественно ожидать, что более детальное рассмотрение указанной аналогии окажется полезным при исследовании временных рядов.

В качестве одного из характерных примеров укажем на известный в спектроскопии эффект - появление вблизи сильных спектральных линий серии спутников (*духов Роуланда*), обусловленных периодической ошибкой шага винта, подающего заготовку при нарезании дифракционной решетки (Малышев, 1979). Если $\nu_{0,m}$ - частота основной линии в спектре m-го порядка, то частота k-го духа Роуланда равна

$$\nu_{k,m}^{(g)} = \nu_{0,m}\left(1 \pm \frac{k}{m}\,\frac{s_0}{P_s}\right), \qquad k,\ m = \pm 1,\ \pm 2, \ldots, \tag{3.7.1}$$

где s_0 - пространственный период дифракционной решетки (расстояние между штрихами) и P_s - шаг винта, равный пространственному периоду решетки-духа. При наземных астрономических наблюдениях какого-либо временного процесса аналогом

s_0 будет чаще всего продолжительность суток, и правильной дифракционной решетке в данном случае отвечает равномерный ряд, а периодической ошибке винта - периодическое же изменение интервалов времени между наблюдениями, обусловленное, например, колебаниями длительности истинных солнечных суток (так называемое *уравнение времени*). Очевидно, последнее обстоятельство должно приводить к появлению на периодограмме духов, аналогичных духам Роуланда.

Подобным образом можно найти аналоги в области временных рядов самому факту появления спектров различных порядков, *духам Лаймана* и другим известным в спектроскопии явлениям. Столь же отчетливо прослеживается связь между построением спектральных оценок в теории временных рядов и формированием изображения при помощи оптических элементов, а также методами статистической радиофизики. Например, описываемое уравнениями (1.6.18), (1.6.20) влияние скважности наблюдений эквивалентно *амплитудной модуляции* сигнала, а эффекты неравномерности распределения моментов наблюдения - *частотной модуляции* (см. Харкевич, 1962). Мы не имеем здесь возможности остановиться на указанных вопросах сколько-нибудь подробно и ограничимся кратким обсуждением двух оптических аналогий.

Временные ряды и спектральный анализ излучения при помощи дифракционной решетки. На рис. 3.7.1 представлена схема образования спектра скалярного колебания $f(t)$ при помощи прозрачной дифракционной решетки G и объектива камеры C.

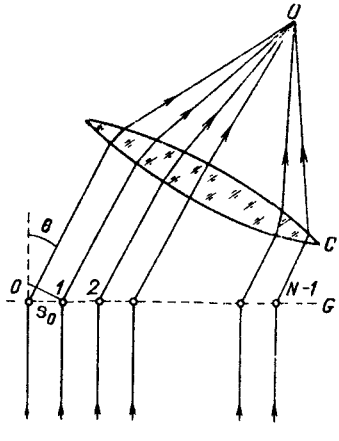


Рис. 3.7.1. Схема образования спектра с помощью дифракционной решетки

Предполагается, что решетка состоит из N штрихов, разделенных равными промежутками шириной s_0. Наглядная интерпретация образования в фокальной плоскости объектива спектров различных порядков $m = 0, \pm 1, \pm 2, \ldots$ при помощи соотношения

$$m\lambda = s_0 \sin\theta \qquad (3.7.2)$$

($\lambda = c/\nu$ - длина волны излучения с частотой ν, c - скорость света) приводится во многих учебниках оптики, и мы не будем ее касаться. Обратимся к эквивалентному подходу, опирающемуся на понятия теории временных рядов.

Поскольку запаздывание колебаний, приходящих в точку наблюдения O от соседних штрихов решетки, равно при фиксированном угле рассеяния

$$\Delta t(\theta) = \frac{s_0}{c} \sin\theta, \qquad (3.7.3)$$

можно представить суммарное колебание, распространяющееся в направлении θ, следующим образом:

$$\varphi(t,\theta) = \sum_{k=0}^{N-1} f\,[t + k\Delta t(\theta)]. \qquad (3.7.4)$$

Подставляя сюда спектральное разложение $f(t)$ в виде (1.3.2) и учитывая определение (1.6.19) спектрального окна данных, получаем

$$\varphi(t,\theta) = N \int_{-\infty}^{\infty} e^{i\cdot 2\pi\nu t}\, W_N(-\nu,\theta)\, F(\nu)\,d\nu, \qquad (3.7.5)$$

где явным образом отмечена зависимость функции W_N от угла рассеяния излучения:

$$W_N(\nu,\theta) = \frac{1}{N} \sum_{k=0}^{N-1} e^{-i\cdot 2\pi\nu k\Delta t(\theta)}. \qquad (3.7.6)$$

Регистрируемый поток излучения $\mathcal{F}(\theta)$ пропорционален усредненному за большой промежуток времени T значению $|\varphi(t)|^2$, т.е.

$$\mathcal{F}(\theta) \propto \frac{1}{T} \int_0^T |\varphi(t)|^2\, dt \propto \int_{-\infty}^{\infty} g(\nu)\,|W_N(\nu,\theta)|^2\, d\nu, \qquad (3.7.7)$$

где мы воспользовались равенством (5) и теоремой Рэлея.

Соотношение (7) показывает, что переход от истинной

спектральной плотности электромагнитных колебаний $g(\nu)$ к наблюдаемому в направлении θ потоку определяется функцией плотности спектрального окна $|W_N|^2$. Как было выяснено в § 2.1, эта функция имеет вид «гребенки», состоящей из единичных «зубцов» шириной $(N\Delta t)^{-1}$, расположенных с шагом Δt^{-1} (см. рис. 2.1.1). В данном случае шаг зависит от угла рассеяния, так что максимумы $|W_N(\nu, \theta)|^2$ расположены на частотах

$$\nu_m = \frac{m}{\Delta t} = m \frac{c}{s_0 \sin \theta}, \quad m = 0, \pm 1, \ldots, \tag{3.7.8}$$

и переход от шкалы частот к длинам волн вновь приводит нас к соотношению (2). При $N \gg 1$ упомянутые максимумы весьма узки, и мы можем приближенно записать (7) в виде

$$\mathcal{I}(\theta) \simeq \text{const} \sum_m g(\nu) \Big|_{\nu = m/\Delta t(\theta)}, \tag{3.7.9}$$

что соответствует наложению в каждой точке наблюдения спектров разных порядков. Заметим, что спектральное разрешение решетки составляет величину порядка $(N\Delta t)^{-1}$, т.е. определяется шириной коллимированного пучка.

Таким образом, *ложные максимумы на периодограмме и спектры различных порядков, построенные дифракционной решеткой, обусловлены одной и той же причиной - регулярным дискретным расположением соответственно моментов наблюдения и штрихов решетки.*

Предположим для простоты, что спектр излучения состоит из единственной линии, расположенной на частоте ν_0. Тогда из (7) следует

$$\mathcal{I}(\theta) \propto |W_N(\nu_0, \theta)|^2 \tag{3.7.10}$$

- в наблюдаемом спектре имеются максимумы при углах рассеяния θ_m, связанных с ν_0 стандартным соотношением $\nu_0 = m/\Delta t(\theta_m)$, $m = 0, \pm 1, \ldots$ Форма спектральных линий полностью определяется видом плотности спектрального окна. Так, при равномерном шаге s_0 решетки контур линии задается функцией

$|\phi_N(\nu\Delta t)|^2$, где $\phi_N(x)$ определена равенством (2.1.4). Если $N \gg 1$, то линия имеет колоколообразный вид с шириной $\delta\nu \sim (N\Delta t)^{-1}$ и незначительными боковыми максимумами (sidelobes) с относительной интенсивностью, не превосходящей 1/23. Ближайший к основному максимум расположен при $\nu \simeq 1{,}5(N\Delta t)^{-1}$. Допустим, однако, что расстояния s_l между штрихами решетки не равны одной и той же величине s_0, т.е. рассмотрим неравномерную решетку. Тогда вместо (6) будем иметь:

$$W_N(\nu,\theta) = \frac{1}{N}\sum_{k=0}^{N-1} \exp\left(i\cdot 2\pi\nu \sum_{l=0}^{k-1}\Delta t_l\right), \qquad (3.7.11)$$

где

$$\Delta t_l = \frac{s_l}{c}\sin\theta, \qquad l = 0, 1, \ldots, N-2. \qquad (3.7.12)$$

Случайное распределение интервалов Δt_l изучалось нами в § 3.2. Предположим теперь, что s_l изменяются периодически с длиной цикла P_s, так что

$$\Delta t_l = \Delta t_0\left[1 + \mu\sin\left(2\pi\frac{l}{L}\right)\right], \qquad (3.7.13)$$

где

$$L = P_s/s_0, \qquad (3.7.14)$$

а параметр μ характеризует глубину модуляции длины интервалов между моментами наблюдения ($|\mu| \leqslant 1$). Полагая $x = \nu\Delta t_0$, имеем согласно (П.1.2):

$$\gamma_k \equiv \frac{t_k}{\Delta t_0} = k + \mu\frac{\sin[(k-1)\pi/L]\,\sin(k\pi/L)}{\sin(\pi/L)}, \qquad (3.7.15)$$

$$|W_N|^2 = \frac{1}{N^2}\left|\sum_{k=0}^{N-1} e^{i\cdot 2\pi x\gamma_k}\right|^2. \qquad (3.7.16)$$

Разумеется, при $\mu = 0$ рассматриваемый временной ряд становится равномерным, а выражение в правой части (16) совпадает с функцией $|\phi_N(x)|^2$, представленной на рис. 2.1.1.

В качестве примера по формулам (15) и (16) при $\mu = 0{,}05$, $L = 10$, $N = 50$ был вычислен спектр скважности, а затем и периодограмма гармонического колебания $y_k = \cos(2\pi\nu_0 t_k)$ при

$\nu_0 = 1/e = 1/2,71828$ (рис. 3.7.2). В спектре скважности вблизи основных линий на частотах $\nu_{0,m} = m/\Delta t_0$ имеются духи, смещенные на величины, кратные $1/L = 0,1$. Именно такой характер спектра и предсказывается формулой (1). Обратим внимание, что интенсивность основных линий быстро убывает, в то время как интенсивность д у́ х о в р а з л и ч н о г о п о р я д к а д о с т и г а е т м а к с и м у м а в р а з н ы х о б л а с т я х с п е к т р а. Общий вид спектра скважности на высоких частотах весьма нерегулярен; как показывают численные примеры, в нем нередко можно видеть максимум значительной интенсивности при $\nu \gg 1$. С возрастанием глубины модуляции $|\mu|$ основные максимумы исчезают уже при $m \simeq 1,2$ и спектр скважности быстро приобретает хаотический вид. Как следствие, еще более запутанными представляются и периодограммы даже простейших колебаний (рис. 3.7.2, б).

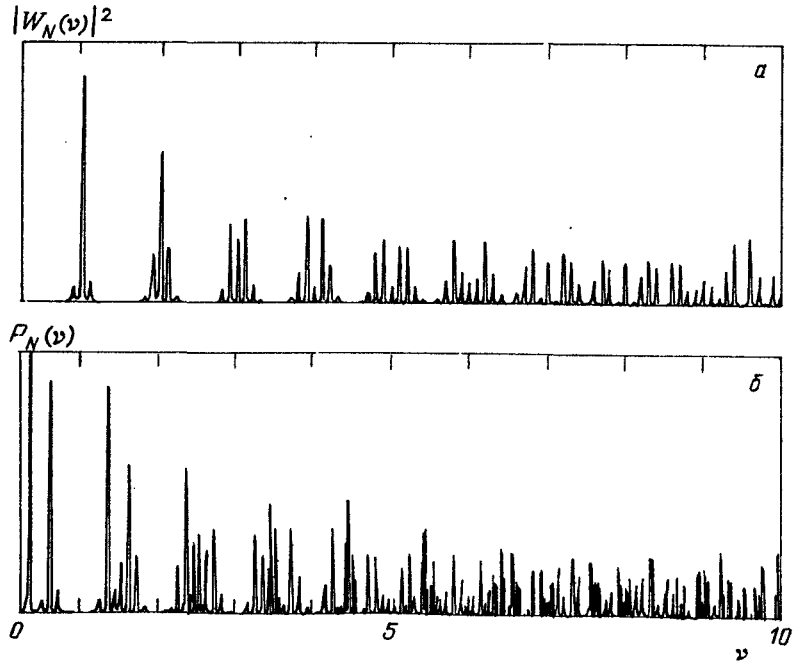


Рис. 3.7.2. Спектр скважности (а) и периодограмма гармонического колебания (б) при периодическом изменении длин интервалов между наблюдениями

Построение пространственного спектра волны с помощью оптических элементов. Основные понятия, используемые при анализе временных процессов, - спектральная плотность, автоковариационная функция, фильтр, весьма наглядно иллюстрируются в теории образования изображения посредством оптических элементов (см., например, Борн и Вольф, 1964; Литвиненко, 1974).

Предположим (рис. 3.7.3), что поле $u(x', y', 0)$ в плоскости S' задано, и требуется найти поле $u(x, y, z)$ в плоскости S после прохождения волны через некоторую оптическую систему L. Для линейной системы переход от S' к S можно представить как линейную фильтрацию с некоторой импульсной характеристикой $h(x, y, z)$:

$$u(x, y, z) = \iint_{-\infty}^{\infty} h(x-x',\ y-y', z)\ u(x', y', 0)\,dx'\,dy'. \qquad (3.7.17)$$

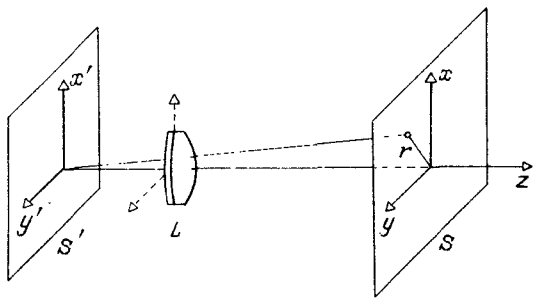


Рис. 3.7.3. Схема формирования изображения оптической системой L

В литературе по оптике $h(x, y, z)$ чаще называют *функцией рассеяния точки* (Point Spread Function - PSF). Переходя в область пространственных частот, имеем по теореме о свертке:

$$U(\sigma_1, \sigma_2, z) = H(\sigma_1, \sigma_2, z)U(\sigma_1, \sigma_2, 0), \qquad (3.7.18)$$

где

$$U(\sigma_1, \sigma_2, z) = \iint_{-\infty}^{\infty} e^{-i\cdot 2\pi(\sigma_1 x+\sigma_2 y)}\, u(x, y, z)\,dx\,dy\ , \qquad (3.7.19)$$

а передаточная функция $H(\sigma_1, \sigma_2, z)$ представляет собой аналогичный двумерный фурье-образ функции рассеяния точки. В дальнейшем мы не будем указывать бесконечные пределы интегрирования.

Итак, поставленная задача может быть полностью решена путем обращения (19), если известна передаточная функция $H(\sigma_1, \sigma_2, z)$. Эта функция является комплексной, так что необходимо знать $|H(\sigma_1, \sigma_2, z)|$ - так называемую *частотно-контрастную характеристику* и $\arg H(\sigma_1, \sigma_2, z)$ - *фазо-контрастную характеристику.*

Пример 3.7.1. Слой пространства. Предположим, что между S' и S нет оптических элементов. Тогда передаточная функция для монохроматической волны частоты $\nu = c/\lambda$ без труда находится из волнового уравнения:

$$H(\sigma_1, \sigma_2, z) = e^{i \cdot 2\pi z \sqrt{\lambda^{-2} - \sigma_1^2 - \sigma_2^2}} . \qquad (3.7.20)$$

Очевидно, $|H| = 1$ при $\sigma \equiv \sqrt{\sigma_1^2 + \sigma_2^2} < 1/\lambda$, т.е. амплитудное пропускание слоя пространства равно 1 в области характерных пространственных размеров $l \equiv 1/\sigma$, превосходящих длину волны λ. Функция рассеяния точки, отвечающая (20), в области $r = \sqrt{x^2 + y^2 + z^2} \gg \lambda$ равна

$$h(x, y, z) = \frac{k}{2\pi i} \frac{z\, e^{ikr}}{r^2}, \quad k = \frac{2\pi}{\lambda} . \qquad (3.7.21)$$

Выражение (21) соответствует пренебрежению квазистационарным полем и сохранению только поля излучения. В параксиальном приближении это просто сферическая волна, распространяющаяся вдоль оси z.

Предположим далее, что поперечные размеры интересующей нас области значительно меньше удаления от источника, т.е. $x^2 + y^2 \ll z^2$. Тогда из (21) следует:

$$h(x, y, z) \approx \frac{k}{2\pi i z} e^{ikz\left(1 + \frac{x^2 + y^2}{2z^2}\right)} , \qquad (3.7.22)$$

что соответствует известному приближению Френеля. Согласно (17) и (22) мы имеем в области приближения Френеля

$$u(x,y,z) = \frac{k}{2\pi i}\ \frac{e^{ikz}}{z}\iint e^{\frac{ik}{2z}\left[(x-x')^2+(y-y')^2\right]} \times$$

$$\times\ u(x',y',0)dx'dy'. \qquad (3.7.23)$$

Дальнейшего упрощения можно достичь, если оставить в показателе степени только члены, линейные по x' и y', так что

$$u(x,y,z) = C\iint e^{-i\left(\frac{kx}{z}x'+\frac{ky}{z}y'\right)} u(x',y',0)dx'dy', \qquad (3.7.24)$$

где

$$C = \frac{k}{2\pi iz}\,e^{ik\left(z+\frac{x^2+y^2}{2z}\right)}. \qquad (3.7.25)$$

Выражение (24) соответствует приближению Фраунгофера. Поскольку двойной интеграл в правой части (24) представляет собой трансформанту Фурье функции $u(x',y',0)$, мы приходим к важному результату:

$$u(x,y,z) = CU\left(\frac{x}{\lambda z},\ \frac{y}{\lambda z},0\right) \qquad (3.7.26)$$

– в приближении Фраунгофера поле на выходе системы пропорционально фурье-образу исходного поля. Множитель C зависит от всех трех координат, но при рассмотрении изображения в плоскости S (т.е. при изменении только x и y) меняется лишь фаза волны. Поэтому интенсивность волны в плоскости S

$$I(x,y,z) = \frac{k^2}{4\pi z^2}\left|U\left(\frac{x}{\lambda z},\ \frac{y}{\lambda z},0\right)\right|^2 \qquad (3.7.27)$$

пропорциональна спектральной плотности входного поля.

Пример 3.7.2. Линза и слой пространства. Пусть тонкая линза L (рис. 3.7.3) расположена непосредственно за плоскостью входного изображения S'. Тогда для нахождения поля в плоскости S необходимо сначала учесть возмущение, вносимое

линзой, а затем использовать приведенные выше соотношения для слоя пространства.

Обозначим через n показатель преломления вещества, из которого изготовлена линза, и через $\Delta z(x',y')$ - толщину линзы в точке с координатами (x',y'). Ввиду различия значений скорости света в воздухе и веществе линзы луч, прошедший через линзу в точке (x',y'), испытывает сдвиг фазы $\Delta\varphi(x',y') = 2\pi(n-1)\Delta z(x',y')/\lambda$. Если линза ограничена параболическими поверхностями с радиусами кривизны при вершинах R_1 и R_2, то

$$\Delta\varphi(x',y') = k(n-1)\left[\text{const}-\left(\frac{1}{R_1}+\frac{1}{R_2}\right)\frac{x'^2+y'^2}{2}\right], \qquad (3.7.28)$$

где $k = 2\pi/\lambda$- волновое число (формула приближенно справедлива и для тонкой линзы со сферическими поверхностями). Прошедшая через линзу волна описывается соотношением

$$u_1(x',y',0) = u(x',y',0)e^{i\Delta\varphi(x,'y')} =$$

$$= u(x',y',0)a\ e^{-ik\frac{x'^2+y'^2}{2f}}, \qquad (3.7.29)$$

где $|a| = 1$ в случае прозрачной линзы и f - фокусное расстояние линзы, определенное формулой

$$\frac{1}{f} = (n-1)\left(\frac{1}{R_1}+\frac{1}{R_2}\right). \qquad (3.7.30)$$

Предположим теперь, что плоскость изображения удалена от S' и линзы на расстояние f. Полагая в (23) $z = f$ и подставляя в это соотношение u_1 из (29) вместо $u(x',y',0)$, получаем

$$u(x,y,f) = \frac{ak}{2\pi if}\ e^{ik\left(f+\frac{x^2+y^2}{2f}\right)}\ U\left(\frac{x}{\lambda f},\ \frac{y}{\lambda f},\ 0\right). \qquad (3.7.31)$$

Таким образом, в приближении Френеля линза и слой пространства реализуют преобразование Фурье входного поля, а интенсивность в плоскости изображения пропорциональна спектральной плотности

входного поля:

$$I(x,y,f) = \text{const}\cdot\left|U\left(\frac{x}{\lambda f},\ \frac{y}{\lambda f},0\right)\right|^2. \qquad (3.7.32)$$

В отличие от соотношения (27), отвечающего приближению Фраунгофера, в (32) спектр строится на расстоянии $f \ll z$ от плоскости входного изображения, что обусловлено, конечно, вносимым линзой фазовым сдвигом.

Если на линзу падает неограниченная плоская волна $u(x',y',0) = u_0 = \text{const}$, то, как показывают (32) и определение δ-функции (П.2.4), изображение имеет точечный вид. Допустим, однако, что линза ограничена круговой диафрагмой диаметром D. При этом условии, соответствующем наблюдению звезды с помощью телескопа, из (32) следует известная формула Эйри (1835):

$$I(x,y,f) = I_0\left[\frac{2J_1(\rho)}{\rho}\right]^2, \qquad (3.7.33)$$

где I_0 - центральная интенсивность, $J_1(\rho)$ - функция Бесселя, $\rho = r/r_0$, $r = \sqrt{x^2 + y^2}$ - расстояние от центра изображения, r_0 - характерный радиус дифракционной картины, равный $\lambda/(\pi A)$, $A = D/f$ - относительное отверстие объектива.

Как и в теории временных процессов, где ограниченность реализации интервалом времени T приводит к изображенному на рис. 1.3.2 профилю спектральной линии вида $\text{sinc}^2(\nu T)$, ограниченность волнового фронта порождает вторичные дифракционные максимумы в изображении. Аналогия станет еще более полной, если в (32) вместо круговой апертуры рассмотреть длинную щель шириной D, расположенную в направлении оси y', а в изображении ограничиться осевым сечением поперек изображения щели. Тогда интенсивность

$$I(x,0) = I_0\left[\frac{\sin(x/x_0)}{x/x_0}\right]^2 = I_0\,\text{sinc}^2\left(\frac{Dx}{\lambda f}\right), \qquad (3.7.34)$$

где $x_0 = \lambda/(\pi A)$, $A = D/f$. Степень выраженности вторичных дифракционных максимумов определяется не столько конечностью волнового фронта, сколько резким его обрывом, поэтому можно ввести апертурную маску с постепенно уменьшающимся к границам пропусканием. Такая операция, называемая в оптике апо-

дизацией, позволяет значительно уменьшить интенсивность вторичных максимумов, но при этом, разумеется, ширина основного пика несколько возрастает (ср. со сглаживанием временного окна, § 2.4). Иногда аподизация позволяет более уверенно различить мелкие детали. Например, при визуальном поиске тесных двойных звезд возможна ситуация, когда яркие вторичные максимумы дифракционной картины мешают разделить компоненты, и удачная аподизация окажется эффективной. Вместе с тем ясно, что таким путем невозможно принципиально повысить разрешение - оно останется порядка λ/D в угловых единицах (см. формулы (33) и (34)).

Существенно повысить разрешение за дифракционный предел удается на пути, реализованном в численном виде при рассмотрении метода максимальной энтропии (§ 2.5). Действительно, мы выяснили, что интенсивность построенного в фокальной плоскости объектива изображения пропорциональна спектральной плотности падающей на объектив волны, ограниченной заданной апертурой. Обратное преобразование Фурье изображения, выполненное численно либо с помощью оптической системы, дает автоковариационную функцию этой волны. Рассматривая известную автоковариационную функцию в качестве ограничения типа (2.5.25) и отыскивая такое ее продолжение, которое придает максимум энтропии (2.5.24), можно более точно восстановить спектр падающего на диафрагму волнового фронта и тем самым повысить разрешение. Как упоминалось в § 2.5, при этом фактически находится оптимальное продолжение известной части волнового фронта за пределы апертуры, что и объясняет эффект сверхразрешения. Другая возможность превзойти дифракционный предел разрешения опирается на аналитическое продолжение волнового фронта (см. Василенко и Тараторин, 1986).

§ 3.8. Некоторые общие методы решения обратных задач

В § 1.2 отмечалось, что анализ временных рядов относится к группе так называемых *обратных задач* математической физики, связанных с восстановлением информации о процессе, искаженной действием измерительного прибора и шума. Такая по-

становка задачи тем более справедлива, если под t понимается пространственная координата и речь идет о восстановлении одномерных изображений (см. § 1.1). Поскольку все предыдущее изложение в равной мере касалось и собственно временных рядов, и пространственных изображений, целесообразно дать здесь хотя бы беглое представление о методах решения обратных задач, которые в настоящее время широко используются в прикладных исследованиях. По необходимости ограничимся рассмотрением одномерных задач для *уравнений типа свертки* (формальное усложнение задачи не изменяет существа дела).

Предположим, что искомый сигнал $s(t)$ поступает на вход «прибора» (рис. 3.8.1), действие которого сводится к однородному размыванию входного процесса, описываемому *аппаратной функцией* $k(\tau)$:

$$f(t) = \int_{-\infty}^{\infty} k(t-t')s(t')dt', \qquad (3.8.1)$$

где для удобства принимается условие

$$\int_{-\infty}^{\infty} k(\tau)d\tau = 1, \qquad (3.8.2)$$

обеспечивающее равенство полных интегралов от входного и размытого изображений. Далее процесс проходит через систему, условно названную «шум»; ее действие заключается в аддитивном внесении случайного компонента $n(t)$, так что на выходе экспериментатор имеет дело с процессом $x(t) = f(t) + n(t)$, или

$$x(t) = \int_{-\infty}^{\infty} k(t-t')\ s(t')\ dt' + n(t). \qquad (3.8.3)$$

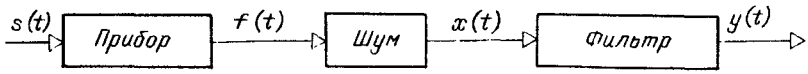


Рис. 3.8.1. Схема преобразования сигнала

Задача ставится следующим образом: по заданной реализации $x(t)$, известным статистическим характеристикам шума и некоторой априорной информации о сигнале восстановить его в определенном смысле наиболее полно.

Указанная формулировка задачи охватывает чрезвычайно широкий круг проблем, возникающих в физике и технике; к ним относятся, в частности, восстановление размытых земной атмосферой изображений астрономических объектов и искусственных космических тел, повышение качества и разрешения записей спектров, выделение слабых сигналов на фоне шума и т.д. Сюда же можно отнести и обсуждавшуюся в § 1.6 проблему восстановления истинного спектра процесса по его наблюдаемому дискретному спектру и заданной скважности наблюдений.

Неустойчивость решения и фильтрация. Рассмотрим сначала идеальный случай, когда шум отсутствует и задача сводится к решению уравнения (1), совпадающего с представлением (1.3.7) функции в виде свертки двух других функций. Как и в (1.3.8) находим:

$$F(\nu) = K(\nu)S(\nu), \tag{3.8.4}$$

где прописными буквами обозначены соответствующие фурье-образы, и далее

$$S(\nu) = \frac{1}{K(\nu)} F(\nu), \tag{3.8.5}$$

$$s(t) = \int_{-\infty}^{\infty} e^{i\cdot 2\pi\nu t} S(\nu)d\nu. \tag{3.8.6}$$

Последние две формулы дают стандартное решение уравнений типа свертки (1).

Казалось бы, не составляет труда применить обращение преобразования Фурье для решения уравнения (3). Мы имеем:

$$X(\nu) = K(\nu)S(\nu)+N(\nu), \tag{3.8.7}$$

откуда

$$S(\nu) = \frac{1}{K(\nu)} [X(\nu)-N(\nu)], \tag{3.8.8}$$

и подстановка (8) в (6) дает формальное решение задачи. Однако известными являются лишь статистические свойства шума, но не конкретная его реализация $n(t)$, вошедшая в уравнение (3). Поэтому в (8) функция $N(\nu)$ остается, строго говоря, неизвестной. Поскольку процесс $n(t)$ предполагается стационарным, можно попытаться заменить $N(\nu)$ фурье-образом некоторой известной реализации шума $\tilde{N}(\nu)$ с теми же статистичес-

кими характеристиками (например, использовать запись шума на участке фотопластинки, свободном от анализируемого изображения). Тогда мы приходим к рассмотрению функции

$$\tilde{S}(\nu) \equiv \frac{1}{K(\nu)} \quad [X(\nu) - \tilde{N}(\nu)] \qquad (3.8.9)$$

в качестве приближения к точному спектру $S(\nu)$ и соответственно $\tilde{s}(t)$ - к искомому сигналу. Как легко видеть,

$$\tilde{s}(t) = s(t) + \int_{-\infty}^{\infty} e^{i\cdot 2\pi\nu t} \quad \frac{N(\nu) - \tilde{N}(\nu)}{K(\nu)} \, d\nu \ . \qquad (3.8.10)$$

Это соотношение ясно показывает причину широко известной на практике неустойчивости решения уравнения (3) относительно малых ошибок результатов измерений или добавки шума. Размывание сигнала при прохождении через «прибор» формально описывается уменьшением эффективности передачи высоких частот, т.е. быстрым убыванием $|K(\nu)|$ при $\nu \to \infty$ (рис. 3.8.2). С другой стороны, шум обычно простирается до частот, далеко выходящих за пределы области спектра, содержащей информацию о сигнале. Поэтому при $\nu \geqslant \nu_*$ (см. рис. 3.8.2), где $|K| \ll |N|$, второе слагаемое в правой части (10) вносит большой по абсолютной величине вклад, не отражающий действительного поведения $S(\nu)$ и, более того, сильно меняющийся в зависимости от реализации шума. Это положение сохранится и в том случае, если под $\tilde{N}(\nu)$ понимать средний спектр многих реализаций (уменьшится лишь дисперсия ошибки $s - \tilde{s}$).

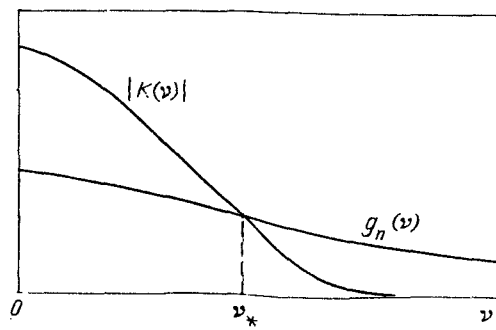


Рис. 3.8.2. Типичные зависимости от частоты спектральной плотности шума $g_n(\nu)$ и модуля передаточной функции прибора $|K(\nu)|$

Таким образом, формальные решения, отвечающие почти сов-

падающим выходным данным, могут отличаться весьма сильно.

Два замечания представляются важными в связи с общей постановкой задачи восстановления изображения и уравнением (3).

Во-первых, процесс размывания изображения имеет вероятностную природу, и лишь при очень большом сигнале можно пренебречь собственными его флуктуациями, т.е. заменить действительный вклад источника его средним значением - первым слагаемым в правой части (3). Другими словами, при написании этого уравнения предполагается бесконечно большое отношение сигнала к внутреннему шуму (в контексте восстановления изображений - радиационному шуму). Если же внутренний шум принимать во внимание, то и при нулевом внешнем шуме возможно лишь неполное решение обратной задачи. Способ максимально правдоподобного восстановления изображения, учитывающий как аддитивный шум, так и естественные флуктуации сигнала, предложен Теребижем (1990).

Во-вторых, мы заметили выше, что принципиальная причина неустойчивости формального решения состоит в отсутствие информации о точном поведении спектра мощности шума в области высоких частот. Но почему замена спектра истинной реализации спектром типичной реализации или средним значением шума не дает удовлетворительных результатов? Суть дела в том же явлении сильных флуктуаций спектра мощности шума, которыми обусловлена несостоятельность спектральных оценок (§ 2.3). Еще А.Шустер (1898) показал, что значения спектра мощности шума, разделенные интервалом частот порядка обратной длительности ряда (или протяженности изображения), независимы друг от друга, а отдельные значения подчиняются экспоненциальному вероятностному закону, т.е. плотность распределения ординаты спектра очень медленно убывает в области больших уклонений от среднего. Поэтому спектр мощности шума (чисто белого или близкого к нему) принципиально имеет «рваный» вид, и спектральные плотности отдельных реализаций сильно отличаются друг от друга. Нетрудно показать, что вероятностные характеристики спектра мощности шума определяют значимость выделения деталей оригинального изображения, и в частности - разделения близко расположенных компонентов.

Простейший выход из положения заключается в построении некоего линейного фильтра с передаточной функцией $H(\nu)$, который устранит из формального решения вклад высокочастотной области спектра $\nu \geq \nu_*$, почти не содержащей информации о сигнале (т.е. сгладит наблюдаемое изображение). В результате фильтрации мы получим на выходе спектр

$$Y(\nu) = H(\nu)X(\nu) = H(\nu)[K(\nu)S(\nu)+N(\nu)], \qquad (3.8.11)$$

а затем и сам выходной процесс

$$y(t) = \int_{-\infty}^{\infty} h(t-t')x(t')dt' , \qquad (3.8.12)$$

который должен быть в определенном смысле близок ко входному сигналу (здесь $h(\tau)$ - импульсная характеристика фильтра, т.е. фурье-образ его передаточной функции). При $N(\nu) \equiv 0$ наилучшим является, как уже говорилось, *инверсный фильтр*

$$H_I(\nu) = \frac{1}{K(\nu)} , \qquad (3.8.13)$$

точно восстанавливающий спектр сигнала $S(\nu)$. В общем случае можно написать

$$H(\nu) = \Phi(\nu) \frac{1}{K(\nu)} , \qquad (3.8.14)$$

где множитель $\Phi(\nu)$ призван устранить неустойчивость, порождаемую высокочастотным «хвостом» шума.

Фильтр Винера строится исходя из условия, что сигнал $y(t)$ на выходе фильтра как можно меньше в среднеквадратическом смысле отличается от входного сигнала $s(t)$, т.е.

$$\varepsilon^2 \equiv \langle [y(t) - s(t)]^2 \rangle = \min \qquad (3.8.15)$$

(А.Н.Колмогоров, 1941 а, б; Н.Винер, 1949). Из (3), (12) и (15) нетрудно показать, что для *винеровского фильтра* (его часто называют также *оптимальным фильтром*) фактор $\Phi(\nu)$ равен

$$\Phi_W(\nu) = \frac{|K(\nu)|^2}{|K(\nu)|^2 + g_n(\nu)/g_s(\nu)} , \qquad (3.8.16)$$

где $g_n(\nu)$ и $g_s(\nu)$ - спектральные плотности шума и сигнала

соответственно. Именно эта априорная информация должна быть задана, чтобы реализовать оптимальный фильтр. При отсутствии размывания изображения $k(\tau) = \delta(\tau)$, $K(\nu) \equiv 1$, и винеровский фильтр описывает восстановление сигнала после аддитивного внесения шума. Разумеется, при отсутствии последнего винеровский фильтр совпадает с инверсным.

Пример 3.8.1. Пусть $n(t)$ - белый шум, так что $g_n(\nu) = g_0 = \text{const}$, а сигнал имеет дисперсию σ^2 и автоковариационную функцию

$$\psi_s(\tau) = \sigma^2 e^{-2\pi\nu_0|\tau|}, \quad -\infty < \tau < \infty, \tag{3.8.17}$$

которой отвечает спектральная плотность

$$g_s(\nu) = \frac{\sigma^2}{\pi\nu_0} \frac{1}{1+(\nu/\nu_0)^2}. \tag{3.8.18}$$

Тогда отношение шум/сигнал в (16) принимает вид

$$\frac{g_n(\nu)}{g_s(\nu)} = \frac{\pi\nu_0 g_0}{\sigma^2} [1+(\nu/\nu_0)^2]. \tag{3.8.19}$$

С ростом частоты это отношение увеличивается, и начиная с некоторого значения ν_* - частоты среза - становится больше $|K(\nu)|^2$. Тогда при $\nu \gg \nu_*$ из (16) следует неравенство $\Phi_W(\nu) \ll 1$, как это и должно быть для фильтра, устраняющего неустойчивый вклад области высоких частот. Фильтр Винера «обрезает» высокие частоты наиболее эффективным в смысле среднеквадратичной метрики образом.

Построим фильтр Винера для практически важного случая, когда аппаратная функция имеет гауссовский профиль:

$$k(\tau) = \frac{1}{\tau_a\sqrt{2\pi}} e^{\frac{\tau^2}{2\tau_a^2}}, \quad K(\nu) = e^{-\frac{(2\pi\nu\tau_a)^2}{2}}, \tag{3.8.20}$$

где τ_a - характерная ширина размытия изображения, обусловленного прибором. Тогда (16) принимает вид

$$\Phi_W(\nu) = [1+\alpha_0(1+x^2)e^{(\beta x)^2}]^{-1}, \tag{3.8.21}$$

где постоянные

$$\alpha_0 = \frac{\pi\nu_0 g_0}{\sigma^2}, \quad \beta = 2\pi\nu_0\tau_a, \tag{3.8.22}$$

а $x = \nu/\nu_0$ – безразмерная частота. В соответствии с (19) α_0 есть отношение шум/сигнал на нулевой частоте, параметр β пропорционален отношению ширин аппаратной и автоковариационной функций. Фактор $\Phi_w(\nu)$ в соответствии с (21) при $\alpha_0 = 0{,}1$, $\beta = 1{,}0$ представлен на рис. 3.8.3.

Гомоморфный и управляемый фильтры. Как показывает практика, фильтр Винера придает слишком большой вес гладкости изображения, так что итоговое разрешение невысоко. Между тем, визуальный анализ изображения, остающийся во многих отношениях недостижимым образцом для машинных алгоритмов, опирается и на информацию в области частот, превосходящих частоту среза винеровского фильтра. Желание расширить исследуемый диапазон частот привело к введению *гомоморфного фильтра*, который определяется условием равенства спектров мощности исходного и восстановленного изображений (см., например, Василенко и Тараторин, 1986).

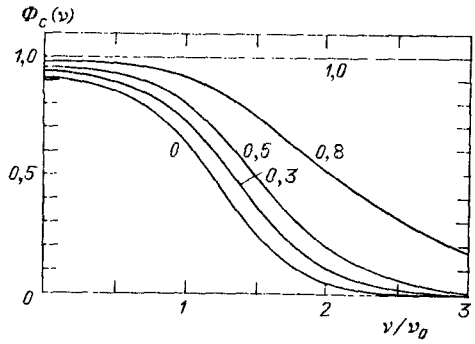


Рис. 3.8.3. Фактор $\phi_c(\nu)$ для управляемого фильтра согласно (25) при $\gamma = 0$ (оптимальный фильтр); 0,3, 0,5, 0,8 и 1,0 (инверсный фильтр)

Пусть $g_s(\nu)$, $g_y(\nu)$ и $g_n(\nu)$ – соответственно спектральные плотности сигнала $s(t)$, выходного процесса $y(t)$ и шума $n(t)$, предполагающегося независимым от сигнала. Накладывая требование $g_y(\nu) = g_s(\nu)$, из (11) и (14) нетрудно получить для гомоморфного фильтра

$$\Phi_H(\nu) = \Phi_w^{1/2}(\nu) = \frac{|K(\nu)|}{\sqrt{|K(\nu)|^2 + g_n(\nu)/g_s(\nu)}}. \qquad (3.8.23)$$

Таким образом, фактор $\Phi(\nu)$ при гомоморфной фильтрации равен

квадратному корню из аналогичной величины для фильтра Винера. Переходя от Φ к частотной характеристике H согласно (14), мы находим, что характеристика гомоморфного фильтра представляет собой среднее геометрическое частотных характеристик инверсного и винеровского фильтров, т.е. $[H_I(\nu)H_W(\nu)]^{1/2}$. Если заменить здесь показатель 1/2 произвольной степенью $\gamma \leqslant 1$, то мы придем к *управляемому фильтру*

$$H_C(\nu) = H_I^{\gamma}(\nu)H_W^{1-\gamma}(\nu). \qquad (3.8.24)$$

Он применяется в тех случаях, когда требуется найти компромисс между гладкостью изображения и разрешением. Из (24) и (14) следует простое соотношение:

$$\Phi_C(\nu) = \Phi_W^{1-\gamma}(\nu), \qquad 0 \leqslant \gamma \leqslant 1. \qquad (3.8.25)$$

Для частного случая, рассматриваемого в примере 3.8.1, это соотношение иллюстрируется рис. 3.8.3.

Регуляризация по Тихонову. Строгое применение фильтра Винера (16) требует знания отношения шума к сигналу g_n/g_s на всех частотах. Аналогичные сведения нужны и при использовании других упомянутых выше фильтров. Сама по себе необходимость иметь априорную информацию совершенно естественна - невозможно восстановить сигнал, ничего не зная о нем и о шуме. Вместе с тем ситуация, когда экспериментатору известны спектральные плотности $g_n(\nu)$ и $g_s(\nu)$, является исключительной; обычно он располагает значительно меньшими сведениями. Поэтому возникает вопрос о построении устойчивых решений уравнения (3) в условиях, когда априорная информация о сигнале скудна. Метод регуляризации А.Н. Тихонова (см. Тихонов и Арсенин, 1979; Гончарский и др., 1985) предназначен для отыскания устойчивых решений широкого класса уравнений именно в таких условиях. В отношении уравнений типа свертки основные результаты регуляризации по Тихонову можно кратко суммировать следующим образом.

1) Выяснены условия, при которых фактор $\Phi(\nu)$ в (14) обеспечивает устойчивость решения (является *регуляризующим*).

2) Доказано, что фактор вида

$$\Phi(\nu,\alpha) = \frac{|K(\nu)|^2}{|K(\nu)|^2 + \alpha M(\nu)} \qquad (3.8.26)$$

при достаточно мягких ограничениях на функцию $M(\nu)$ относится к классу регуляризующих (здесь α - некоторый параметр).

3) Среди множителей вида (26) наибольшую близость приближенного и точного решений уравнения (3) в смысле среднеквадратического отклонения обеспечивает фильтр Винера.

4) Используя меньшую априорную информацию по сравнению с необходимой для построения фильтра Винера, можно заменить отношение $g_n(\nu)/g_s(\nu)$ близкой к нему функцией, например, принять

$$M(\nu) = 1 + (\nu/\nu_0)^2, \qquad (3.8.27)$$

что совпадает с выбранным в примере 3.8.1 представлением (19).

Таким образом, нахождение регуляризованного решения уравнения типа свертки сводится к последовательному вычислению того значения параметра α, которое соответствует заданной неточности экспериментальных данных (дисперсии шума). С обоснованием метода регуляризации и его приложениями к практическим задачам можно познакомиться по цитированным выше книгам, а также книге Ватутина и др. (1985).

Нелинейные методы восстановления изображения. Все перечисленные выше алгоритмы относятся к группе линейных методов. При относительной простоте вычислений линейным методам присущи принципиальные недостатки значительно снижающие эффективность восстановления (см., например, Сороко, 1981; Демченко и Курчаков, 1983; Василенко и Тараторин, 1986; Фриден, 1979).

Прежде всего, линейные методы не учитывают непосредственным образом факт *неотрицательности* решения. Это, казалось бы, простое ограничение оказывается чрезвычайно важным для качества восстановленного изображения. Дополнительное введение указанного ограничения неоднозначно и делает процедуру нелинейной. Далее, линейные методы в принципе не могут восстановить информацию об изображении, которая содержится в области частот, *превосходящих* частоту среза аппаратной функции, т.е. они не позволяют получить высокое

разрешение в изображении. Наконец, вся процедура выбора восстанавливающего линейного алгоритма представляется в значительной мере *искусственной*.

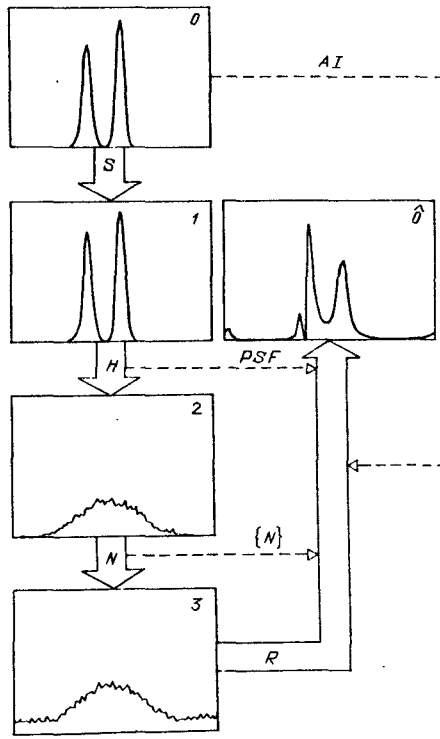


Рис. 3.8.4. Схема формирования и восстановления изображения. Использованы обозначения: o - оригинал; S - образование конечной выборки; 1 - реализация оригинала ; H - система формирования изображения; 2 - размытое изображение с флуктуациями; N - аддитивный шум; 3 - конечное изображение; R - алгоритм восстановления; $\hat{o}$ - оценка оригинала; AI - априорная информация; PSF - функция рассеяния точки (аппаратная функция); *{N}* - средние характеристики шума

Нелинейные информационные методы восстановления изображения устойчивы к шумам, естественным образом учитывают неотрицательность оригинала и позволяют получить более высокое разрешение, чем этого можно ожидать, исходя из простой оценки ширины аппаратной функции. *Эффект сверхразрешения* обусловлен оптимальной экстраполяцией наблюдаемого изобра-

жения, допустимой при данном отношении сигнала к шуму и данной априорной информации (см. § 2.5).

Вместе с тем следует сказать, что обоснование нелинейных методов не достигло пока надлежащей ясности. Фактически каждый из них вводит предположения *ad hoc*, не вытекающие с неизбежностью из существа дела. Само наличие многих нелинейных методов, а к ним относятся несколько вариантов метода максимума энтропии, разновидности метода аналитического продолжения спектра, максимума информации и пр., показывает отсутствие единого подхода, причем такое положение не диктуется разнообразием априорной информации. В частности, сейчас выясняется, что требование максимальной энтропии решения, ограниченной лишь «связями» вида (3), не гарантирует решения, близкого к наиболее вероятному (Фриден, 1985). По-видимому, логически безупречным можно признать лишь интуитивно очевидный подход, основанный *на методе максимума правдоподобия* (ММП). Он требует, чтобы в качестве оценки оригинала было принято такое распределение яркости, при котором наиболее вероятно получить реально наблюдавшееся изображение. Последовательное применение ММП к задаче восстановления изображений дано в работе Теребижа (1990). Рис. 3.8.4 включает действительный пример восстановления при помощи предложенного в этой работе метода.

Нелинейные информационные методы показали столь высокую эффективность, что стали основными при анализе и восстановлении изображений, тем более что быстро возрастающая мощность компьютеров нивелирует требования, предъявляемые линейными и нелинейными алгоритмами.

ГЛАВА 4

НЕПАРАМЕТРИЧЕСКИЕ МЕТОДЫ ВЫДЕЛЕНИЯ СКРЫТОЙ ПЕРИОДИЧНОСТИ

Предыдущие главы были посвящены главным образом нахождению эффективных оценок истинной спектральной плотности временного процесса $g(\nu)$ по результатам N дискретных измерений. Было выяснено, в частности, что исходной величиной для получения таких оценок может служить выборочная спектральная плотность $g_N(\nu) = N\Delta t P_N(\nu)$, где $P_N(\nu)$ - периодограмма Шустера (2.1.22). Для равномерных временных рядов удается существенно повысить надежность получаемой оценки $g(\nu)$ в рамках метода максимума энтропии и других связанных с ним методов. В случае неравномерных рядов оказываются полезными соответствующие обобщения периодограммы типа (1.6.12) и (3.1.17).

Изложенный подход естественным образом опирается на дискретное преобразование Фурье данных наблюдений. Между тем, если ограничиться только задачей выделения скрытой периодичности, т.е. линейчатого спектра процесса, то анализ Фурье не обязательно наиболее эффективен. Один из примеров к этому утверждению - случай низкого отношения сигнала к шуму при сильно отличающейся от синусоидальной форме колебания - приведен в § 1.7. (Этот параграф следует рассматривать как вводный к настоящей главе). Там же было отмечено, что периодограмма Шустера характеризует согласованность распределения точек на фазовой диаграмме по отношению к заданной наперед форме колебания - синусоиде. Отказавшись от сравнения именно с гармоническим колебанием и характеризуя только степень *внутренней* упорядоченности распределения точек на фазовой диаграмме, мы приходим к группе методов,

называемых *непараметрическими* или *методами сворачивания эпох* (epoch folding). Первый из этих терминов представляется более точным, поскольку почти все методы, в том числе и анализ Фурье, используют предварительное сворачивание ряда с набором пробных периодов.

Как уже отмечалось в § 1.7, мера внутренней упорядоченности фазовой диаграммы $\Gamma(\nu)$ для какой-либо пробной частоты ν может быть определена различными способами, так что имеется целый ряд непараметрических методов, отличающихся лишь выбором $\Gamma(\nu)$. Часть из них описана в последующих параграфах данной главы. Выбор метода для изучения конкретного ряда определяется многими обстоятельствами: отношением сигнала к шуму, характером скважности данных, наличием предварительных сведений и т.д. Вместе с тем следует иметь в виду, что на практике применение различных методов к одному и тому же временному ряду обычно дает взаимно дополнительную информацию, и потому разумно не противопоставлять отличающиеся подходы, а использовать их совместно для более глубокого описания изучаемого процесса.

§ 4.1. Статистика Аббе-Лафлера-Кинмана

Рассмотрим временной ряд $\{t_k, y_k\}$, $k = 0, 1, \ldots, N-1$, включающий N не обязательно равноотстоящих отсчетов процесса $y(t)$. Для каждого момента времени t_k при заданном пробном периоде $P = \nu^{-1}$ вычислим фазу

$$x_k(\nu) = \mathrm{FR}(\nu(t_k - t_*)], \quad k = 0, 1, \ldots, N-1. \qquad (4.1.1)$$

где $\mathrm{FR}(z)$ - дробная часть z, t_* - произвольный момент начала отсчета фазы. Рассмотрим далее все фазовые диаграммы, т.е. зависимости y_k от $x_k(\nu)$, для набора пробных частот с равномерным шагом по частоте (1.6.33), обеспечивающим должное разрешение.

Описываемая в данном параграфе характеристика степени разброса точек на фазовой диаграмме строится следующим образом. Расположим фазы x_k в порядке возрастания и обозначим упорядоченные значения посредством $x_{[k]}$, так что

$$0 \leqslant x_{[0]} \leqslant x_{[1]} \leqslant \ldots \leqslant x_{[N-1]} < 1. \qquad (4.1.2)$$

Пусть

$$\vec{y}_k \equiv y(x_{[k]}) \qquad (4.1.3)$$

- соответствующие (2) отсчеты ряда, последовательные не на оси времени, а на фазовой диаграмме (т.е. близкие в смысле пробного периода P). Ясно, что величина

$$s^2_{in} = \frac{1}{2(N-1)} \sum_{k=0}^{N-2} (\vec{y}_{k+1} - \vec{y}_k)^2 \qquad (4.1.4)$$

характеризует упорядоченность точек на фазовой диаграмме: если последовательные точки в среднем близки друг к другу, то s^2_{in} мала, а если они расположены совершенно хаотично, то s^2_{in} будет величиной порядка общей дисперсии ряда

$$s^2 = \frac{1}{N-1} \sum_{k=0}^{N-1} (y_k - \bar{y})^2 \qquad (4.1.5)$$

(см. рис. 1.7.2). Таким образом, следует ожидать, что отношение

$$\theta_N(\nu) = \frac{\sum_{k=0}^{N-2} \left(\vec{y}_{k+1} - \vec{y}_k\right)^2}{2\sum_{k=0}^{N-1} \left(y_k - \bar{y}\right)^2} \qquad (4.1.6)$$

будет существенно меньше единицы при значении пробного периода, близком к истинному периоду P_0, и $\theta_N \sim 1$ при других частотах (рис. 1.7.1). В первом случае точки на фазовой диаграмме будут располагаться тесно вокруг некоторой средней кривой, которая и отражает форму искомой периодичности.

Выражение (6) представляет собой современную форму статистики *критерия Аббе*, предназначенного для выделения тренда (Кроу и др., 1960; Большев и Смирнов, 1965). Аналогичный параметр A_N был введен с этой целью в (3.5.18), где рассматриваются отсчеты, последовательные во времени. Применительно к фазовой диаграмме статистика Аббе была предложена Лафлером и Кинманом (1965), которые вычисляли фактически величину $2\theta_N(\nu)$. Поскольку теоретические исследования и практические приложения критерия Аббе проводились задолго до появления указанной астрономической работы, мы сохраним фактор 2 в знаменателе (6), а саму величину $\theta_N(\nu)$, опреде-

ленную этой формулой, будем называть *статистикой Аббе-Лафлера-Кинмана* (ALK).

Как и при выяснении значимости максимумов периодограммы Шустера в § 2.3, оценка значимости минимумов $\theta_N(\nu)$ должна опираться на распределение статистики ALK для совокупности независимых, одинаково распределенных величин $\vec{y}_k$ (нулевая гипотеза). Начиная с Дж. фон Неймана (1941), функция распределения θ_N для случая нормальных $\vec{y}_k$ изучалась многими авторами (см. ссылки в работе Хека и др., 1985). Распределение θ_N сосредоточено в интервале $[1-\cos(\pi/N), 1+\cos(\pi/N)]$, среднее значение и дисперсия этой величины равны

$$\langle\theta_N\rangle = 1, \qquad (4.1.7)$$

$$\sigma^2(\theta_N) = \frac{N-2}{N^2-1}. \qquad (4.1.8)$$

При заданном уровне значимости q отклонение $\theta_N(\nu)$ от 1 считается существенным, если θ_N оказывается меньше критического значения $A_{N,q}$, вычисленного в соответствии с функцией распределения θ_N при нулевой гипотезе. В табл. П.6.1 воспроизводятся результаты вычислений Харта (1941 a, b), заимствованные нами из книги Большева и Смирнова (1965). При $N \geqslant 50$ можно воспользоваться тем обстоятельством, что величина

$$-(1-\theta_N)\sqrt{\frac{2N+1}{2-(1-\theta_N)^2}}. \qquad (4.1.9)$$

распределена приближенно нормально с параметрами (0, 1). Как видно из (8), ширина «шумовой дорожки» периодограммы ALK при $N \gg 1$ составляет величину порядка $1/\sqrt{N}$.

Предположение о нормальности $\{y_k\}$, определяющее возможность теоретического анализа, не является обязательным. Современные компьютеры позволяют численно реализовать метод рандомизации Р. Фишера (Брэдли, 1968; Диаконис и Эфрон, 1983; Л. Немек и Дж. Немек, 1985). Именно, если предположить, что периодичность отсутствует (нулевая гипотеза), то наблюдаемую последовательность $(y_0, y_1, \ldots, y_{N-1})$ можно считать одним из $N!$ возможных, статистически эквивалентных наборов этих же отсчетов, произвольно переставленных между собой.

Периодограммы, вычисленные для части из этих наборов, дадут представление о значимости наблюдаемых отклонений.

Упорядочение точек по фазе согласно (2) и (3) следует производить методом адресной сортировки, при котором время выполнения операций пропорционально числу точек N (см. Эберт и Эдерер, 1985; Пресс и др., 1986).

Полезно указать на тесную связь статистики ALK с параметром, имеющим простую интерпретацию, а именно, коэффициентом автокорреляции последовательности отсчетов на фазовой диаграмме

$$r_N = \frac{1}{(N-1)s^2} \sum_{k=0}^{N-2} (\vec{y}_{k+1} - \bar{y})(\vec{y}_k - \bar{y}) \quad . \qquad (4.1.10)$$

Вычтем из $\vec{y}_{k+1}$ и $\vec{y}_k$ в числителе (6) среднее значение $\bar{y}$, произведем фактическое возведение во вторую степень и пренебрежем различием конечных слагаемых в числителе и знаменателе, что справедливо при $N \gg 1$. Тогда получим искомое соотношение:

$$\theta_N \simeq 1 - r_N \quad . \qquad (4.1.11)$$

Оно выполняется с достаточной точностью уже при $N \geq 20$. Очевидно, что при хаотическом расположении точек будет $|r_N| \ll 1$ и $\theta_N \sim 1$, а при отчетливо выраженной кривой блеска r_N близок к 1 и $\theta_N \ll 1$.

Свойства r_N хорошо изучены (см. Большев и Смирнов, 1965; Хеннан, 1970), так что соотношением (11) можно пользоваться для оценки значимости флуктуаций статистики ALK. Однако можно и сразу вычислить фазовый коэффициент корреляции (10) для набора пробных частот, рассматривая максимумы $r_N(\nu)$ в качестве признаков наличия периодичности. Это тем более удобно, что коэффициент корреляции - привычная величина, значимость которой легко оценивается интуитивно или при помощи широко распространенных таблиц.

§ 4.2. Фазовая группировка данных (методы Уиттекера - Г.Робинсона, Юркевича, Уорнера - Э.Робинсона и Стеллингвефа).

Для вычисления статистики ALK необходимо при каждом зна-

чении пробного периода произвести поточечное упорядочение отсчетов в соответствии с их фазами. Способы адресной сортировки позволяют выполнить эту операцию достаточно быстро, однако можно и вообще избежать полной сортировки, разбивая предварительно весь круговой интервал изменения фазы [0, 1) на подынтервалы (бины) и формируя при каждом значении пробного периода гистограммы отсчетов. Далее на основании таких сгруппированных по бинам данных нужно оценить, насколько хорошо выделяется средняя кривая блеска.

В астрономической литературе представлена группа методов, отличающихся друг от друга способами такой оценки - это методы Уиттекера и Г. Робинсона (1924), Юркевича (1971), Уорнера и Э. Робинсона (1972), Стеллингвефа (1978). Как мы увидим ниже, первые три метода строго эквивалентны между собой, а последний совпадает с ними в основных чертах. Достоинства методов, опирающихся на фазовую группировку данных, связаны с относительно высокой скоростью расчетов и более гладкой, чем для статистики ALK, периодограммой. Вместе с тем указанные методы пригодны лишь при относительно большом количестве измерений и вносят в окончательные результаты субъективность, обусловленную выбором биновой структуры. Кроме того, при современном быстродействии компьютеров выгоды, доставляемые формированием гистограммы вместо сортировки отсчетов, не имеют решающего значения. Приведенные соображения обусловили преимущественное использование метода фазовой группировки на предварительном, поисковом этапе иссследования.

Дисперсионный анализ Р.Фишера. Основой для сопоставления перечисленных выше вариантов метода фазовой группировки является простое тождество, выполняющееся при разделении отсчетов по бинам. Рассмотрим произвольный временной ряд $\{t_k, y_k\}$, $k = 0, \ldots, N-1$, и вычислим для каждого отсчета фазу $x_k(\nu)$ согласно (1.4.1). Разобьем интервал изменения фазы [0,1) на M неперекрывающихся бинов, содержащих последовательно $N_m (m = 1, 2, \ldots, M)$ отсчетов каждый, так что

$$\sum_{m=1}^{M} N_m = N. \qquad (4.2.1)$$

Пусть Y_{mj} обозначает j-й отсчет в m-ом бине ($j = 1, 2, \ldots, N_m$). Введем оценки групповых среднего и дисперсии отсчетов:

$$\bar{Y}_m = \frac{1}{N_m} \sum_{j=1}^{N_m} Y_{mj}, \qquad (4.2.2)$$

$$s_m^2 = \frac{1}{N_m - 1} \sum_{j=1}^{N_m} (Y_{mj} - \bar{Y}_m)^2 ,$$

а также общего среднего значения выборки и полной дисперсии:

$$\bar{Y} = \frac{1}{N} \sum_{k=1}^{N} Y_k = \frac{1}{N} \sum_{m=1}^{M} N_m \bar{Y}_m = \frac{1}{N} \sum_{m=1}^{M} \sum_{j=1}^{N_m} Y_{mj} , \qquad (4.2.3)$$

$$s^2 = \frac{1}{N-1} \sum_{k=1}^{N} (Y_k - \bar{Y})^2 = \frac{1}{N-1} \sum_{m=1}^{M} \sum_{j=1}^{N_m} (Y_{mj} - \bar{Y})^2 .$$

Обозначив вошедшие сюда двойные суммы просто как $\sum_{m,j}$, запишем упомянутое тождество в виде (Фишер, 1937, 1941; Крамер, 1946)

$$\sum_{m,j} (Y_{mj} - \bar{Y})^2 = \sum_{m,j} (\bar{Y}_m - \bar{Y})^2 + \sum_{m,j} (Y_{mj} - \bar{Y}_m)^2 . \qquad (4.2.4)$$

Для доказательства следует в скобках исходной суммы добавить и вычесть $\bar{Y}_m$, возвести во вторую степень и учесть, что сумма перекрестных членов равна нулю ввиду (2).

Таким образом, сумма квадратов всех отклонений от общего среднего разбивается на сумму квадратов отклонений локальных средних от общего среднего и на сумму квадратов отклонений отсчетов от локальных средних внутри бинов. Каждая из этих частичных сумм может рассматриваться как мера «качества» средней кривой блеска на фазовой диаграмме; отличие фазовой картины от хаотической проявляется тем более четко, чем больше первая сумма (Уиттекер и Г. Робинсон, Уорнер и Э. Робинсон) в правой части (4) или меньше вторая сумма (Юркевич, Стеллингвеф). Но ввиду того, что сумма в левой части тождества, т.е. общий разброс данных, остается фиксированной, оба приведенных критерия эквивалентны. В отноше-

нии статистик Уорнера-Э.Робинсона и Юркевича этот факт был отмечен Ковачем (1980).

Принимая во внимание определения (2) и (3), перепишем тождество (4) в следующем виде:

$$(N-1)s^2 = \sum_{m=1}^{M} N_m(\bar{y}_m - \bar{y})^2 + \sum_{m=1}^{M} (N_m - 1)\, s_m^2 . \qquad (4.2.5)$$

Пусть далее

$$v_1^2 = \frac{1}{M-1} \sum_{m=1}^{M} N_m(\bar{y}_m - \bar{y})^2 \qquad (4.2.6)$$

и

$$v_2^2 = \frac{1}{N-M} \sum_{m=1}^{M} (N_m - 1)s_m^2 \qquad (4.2.7)$$

- соответственно характеристики разброса локальных средних и внутренней дисперсии совокупности точек на фазовой диаграмме. Тогда тождество Фишера принимает окончательный вид:

$$\frac{M-1}{N-1} \frac{v_1^2}{s^2} + \frac{N-M}{N-1} \frac{v_2^2}{s^2} = 1. \qquad (4.2.8)$$

Для оценки статистической значимости различных критериев приведем некоторые результаты для случая, когда все отсчеты независимы и нормальны с дисперсией σ^2 (Крамер, 1946). Величины $(N-1)s^2/\sigma^2$, $(M-1)v_1^2/\sigma^2$ и $(N-M)v_2^2/\sigma^2$ подчиняются χ^2-распределению с $N-1$, $M-1$ и $N-M$ степенями свободы соответственно, причем

$$\langle s^2 \rangle = \langle v_1^2 \rangle = \langle v_2^2 \rangle = \sigma^2, \qquad (4.2.9)$$

так называемое *дисперсионное отношение* v_1^2/v_2^2 является чувствительным показателем отклонения от хаотичности: чем теснее фазовое распределение сжимается около некоторой средней кривой, тем оно больше. Фишер (1924, 1941) в явном виде нашел плотность распределения случайной величины z, связанной с дисперсионным отношением формулой

$$e^{2z} = v_1^2/v_2^2. \qquad (4.2.10)$$

Среднее значение и дисперсия отношения (10) равны:

$$\left\langle \frac{v_1^2}{v_2^2} \right\rangle = \frac{N-M}{N-M-2}, \qquad N-M > 2,$$

(4.2.11)

$$D\left(\frac{v_1^2}{v_2^2}\right) = \frac{2(N-M)^2(N-3)}{(M-1)(N-M-2)^2(N-M-4)}, \qquad N-M > 4.$$

При заданном уровне значимости q отклонение от хаотичности считается существенным, если z превосходит критическое значение z_q, вычисленое для распределения Фишера. Таблицы критических значений приведены во многих руководствах (см., например, Большев и Смирнов, 1965).

Как уже отмечалось в предыдущем параграфе, современные компьютеры позволяют реализовать непараметрический метод рандомизации Фишера для проверки значимости наблюдаемых флуктуаций, не требующий предположения о нормальности отсчетов.

Варианты дисперсионного анализа. Исторически первым таким вариантом, предназначенным для выделения скрытой периодичности, явился метод Уиттекера и Г.Робинсона (1924, с.345). Указанные авторы ввели параметр

$$\theta_N^{(WHR)} = \frac{v_1}{s}, \qquad (4.2.12)$$

где v_1 и s определены соотношениями (6) и (3). Возможным периодичностям отвечают максимумы зависимости $\theta_N^{(WHR)}$ от частоты, названной Уиттекером и Г.Робинсоном *периодограммой* по аналогии с определением, данным Шустером (1898).

Юркевич (1971) рассматривал в качестве меры упорядоченности точек на фазовой диаграмме величину

$$\theta_N^{(J)} = (N - M)\, v_2^2, \qquad (4.2.13)$$

минимумы которой соответствуют периодичностям. Целесообразнее нормировать (13) путем деления на s^2 или v_1^2. Чтобы выяснить роль выбора фазовой структуры, расчеты обычно проводят при разных значениях M, скажем, M = 5, 7 и 11. Ковач (1980) полагает, что оптимальное значение M для сигналов разной формы и различного отношения сигнала к шуму находится между 5 и 9. Некоторые усовершенствования методики предложили Маррако и Муццио (1980), связь с периодограммой Шустера обсуждается Пельтом (1980), Свинглером (1985), программа вычислений приводится дю Пюи и Хоффманом (1985).

Уорнер и Э. Робинсон (1972) применили близкую к (12) статистику

$$\theta_N^{(WR)} = \frac{M-1}{N} v_1^2 . \qquad (4.2.14)$$

Наконец, Стеллингвеф (1978) допускает возможность перекрытия бинов, так что одна и та же точка может входить в несколько групп. При этом условие (1) не соблюдается, а вместо второго из соотношений (7) получаем

$$v_2^2 = \frac{1}{\sum_{m=1}^{M} N_m - M} \sum_{m=1}^{M} (N_m - 1) s_m^2 . \qquad (4.2.15)$$

Эта величина уже не подчиняется χ^2-распределению ввиду зависимости отдельных слагаемых между собой. Применяемый Стеллингвефом критерий основан на отношении

$$\theta_N^{(S)} = \frac{v_2^2}{s^2} , \qquad (4.2.16)$$

сходном с параметром Юркевича (13). Заметим, что статистические выводы Стеллингвефа относительно распределения величины (16) не вполне корректны (Хорман, 1980; Хек и др., 1985). Введением скользящей биновой структуры Стеллингвеф частично избавляется от искусственности, обусловленной группированием данных.

Связь всех приведенных вариантов между собой и с классическим дисперсионным анализом Фишера нетрудно усмотреть из соотношения (8).

Заключительные замечания. Характерная особенность непараметрических методов состоит в том, что они не отдают предпочтения «простым» кривым блеска, как это сделал бы человек при сравнении фазовых диаграмм. Предположим, например, что мы исходим из критерия максимальности дисперсии средних (6). Тогда и последовательность $\{\bar{y}_m\}$, очерчивающая простую кривую с двумя экстремумами, и произвольно переставленная последовательность тех же групповых средних, очерчивающая, скажем, кривую с четырьмя экстремумами, дадут одно и то же значение v_1^2. Формально задача решается путем введения дополнительного критерия «простоты» типа ограниче-

ния количества экстремумов. Более полный анализ должен привлекать методы теории распознавания образов.

§ 4.3. Методы, сохраняющие эффективность при малом количестве измерений («натянутой струны», Рексона и Пельта)

В рассмотренных выше вариантах метода сворачивания эпох мера разброса точек на фазовой диаграмме учитывала лишь отличия их ординат. При значительном количестве наблюдений $N \gg 1$) этого нередко достаточно; нужно лишь, чтобы отсчеты были приблизительно равномерно распределены по фазе. Тогда разница фаз последовательных точек мала, и если они хорошо обрисовывают кривую блеска, то их ординаты должны быть близкими.

При небольших значениях N соседние точки на фазовой диаграмме могут сильно отличаться по фазе, разность их ординат может быть по модулю велика, но это вовсе не свидетельствует с необходимостью об отсутствии подходящей кривой блеска. Предположим, например, что имеется не более 10 измерений чисто периодического процесса без шума. Ясно, что статистики типа ALK при этих условиях не обнаружат резкого экстремума на истинной частоте, т.е. они оказываются неэффективными. Если форма кривой блеска мало отличается от синусоидальной, то здесь явное предпочтение получит анализ Фурье, сравнивающий распределение точек с пробными гармониками. Однако при сложной форме колебания все изложенные методы бессильны выделить периодичность.

Мы приходим к выводу, что мера рассеяния точек на фазовой диаграмме, пригодная и при небольших значениях N, должна включать в себя информацию не только о распределении последовательных отсчетов по величине, но и об их *фазовом удалении*. Анализ равномерного ряда небольшой длины следует проводить при помощи информационных методов, в частности, метода максимума энтропии (§ 2.5), однако для наиболее часто встречающегося случая, когда ряд неравномерен, приходится обращаться к различным вариантам непараметрических критериев.

Разумеется, описываемые ниже методы, работающие при небольших значениях N, будут эффективны и при $N \gg 1$; в пос-

леднем случае можно воспользоваться и более простыми алгоритмами.

Метод «натянутой струны» (string-length) был предложен Э.Бэрком, Ролландом и Боем (1970), работавшими совместно с математиком Сомервиллем (W.H.Somerville). В этом методе указанием на скрытую периодичность является малость величины

$$L_N(\nu) = \sum_{k=0}^{N-1} \sqrt{(\vec{y}^*_{k+1} - \vec{y}^*_k)^2 + (x_{k+1} - x_k)^2} , \qquad (4.3.1)$$

где $\vec{y}^*_k$ (ν) - упорядоченные при заданной пробной частоте по фазе и подходящим образом нормированные отсчеты ряда, замкнутые следующим образом:

$$x_N \equiv x_0 + 1 , \quad \vec{y}^*_N \equiv \vec{y}^*_0 . \qquad (4.3.2)$$

Очевидно, $L_N(\nu)$ равна сумме длин всех отрезков, соединяющих последовательные точки на фазовой диаграмме (из всех распределений выделяется то, для которого «натянутая» на фазовые точки струна имеет наименьшую длину).

Нормировку наблюдаемых отсчетов $\vec{y}_k$, сводящую их к безразмерным величинам $\vec{y}^*_k$, можно осуществить различными способами. Желательно, чтобы разности ординат и фаз входили в (1) с одинаковым весом. Рассматривая простые колебания типа синусоидальных, пилообразных и вариаций лучевой скорости при кеплеровском движении, Дворецкий (1983) предложил следующую нормировку:

$$\vec{y}^*_k = \frac{\vec{y}_k - y_{min}}{2(y_{max} - y_{min})} - \frac{1}{4} . \qquad (4.3.3)$$

При этом значении $\vec{y}^*_k$ заключены в пределах [-1/4, 1/4], а значения параметра L_N для упомянутых простых видов колебания изменяются приблизительно от 1,4 до 1,8. Возможны и другие нормировки, например, типа

$$\vec{y}^*_k = \frac{\vec{y}_k - \bar{y}}{\gamma_N s} , \qquad (4.3.4)$$

где s - стандартное отклонение ординат отсчетов (4.1.5), а γ_N - некоторый параметр порядка единицы, выбираемый для

уравновешивания вклада фаз в (1).

Полезным сопутствующим свойством рассматриваемого метода является более четкая выделенность колебания с основным периодом P_0 по сравнению с колебаниями, включающими несколько таких периодов. Последние имеют сложный вид на фазовой диаграмме, но статистики типа ALK учитывают лишь разброс ординат и выделяют кратные периоды в равной степени с основным.

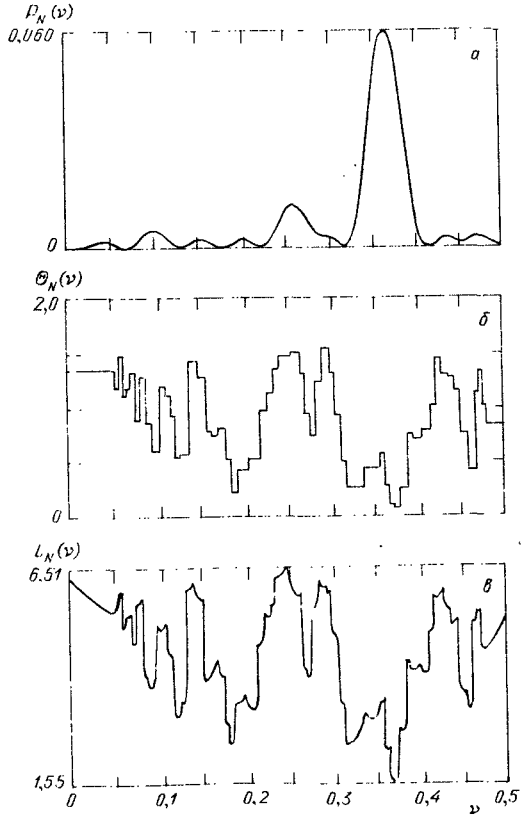


Рис. 4.3.1. Периодограммы Шустера (а), Аббе-Лафлера-Кинмана (б) и «натянутой струны» (в) начального отрезка ряда А, включающего 20 отсчетов $\{f_k\}$

На рис. 4.3.1 приведены периодограммы Шустера, Аббе-Лафлера-Кинмана и метода «натянутой струны» для начального отрезка ряда А, включающего лишь 20 точек и не искаженного шумом (значения $\{f_k\}$, см. § 1.1 и табл. П.7.1). Напомним,

что $\{f_k\}$ представляют собой отсчеты периодического колебания с периодом $P_0 = e \simeq 2,71828$, изображенного на рис. 1.1.5. Все три периодограммы обнаруживают максимальный отклик на частоте периодичности $\nu_0 \simeq 0,368$, однако если $P_N(\nu)$ содержит дополнительные экстремумы лишь на частотах, сопряженных гармоникам $(-2\nu_0 + 1 \simeq 0,264$, $3\nu_0 - 1 \simeq 0,104)$, то $\Theta_N(\nu)$ и $L_N(\nu)$ включают множество минимумов, отвечающих субгармоникам основного периода и сопряженных ему периодов (см. формулу (4.5.3) и обсуждение в § 4.5). Как и следовало ожидать, эти минимумы несколько менее контрастны на $L_N(\nu)$; все же метод «натянутой струны» не позволяет добиться существенного подавления субгармоник. Нетрудно рассчитать, что для чисто синусоидального колебания двойному периоду соответствует значение L_N, превосходящее минимальное приблизительно в 1,57 раза.

Интересно сравнить устойчивость обсуждаемых методов по отношению к введению аддитивного шума. Периодограмма $P_n(\nu)$ первых 20 отсчетов шума $\{\xi_k\}$ ряда A представлена на рис. 4.3.2; периодограммы P_N, Θ_N и L_N значений $\{y_k = f_k + \xi_k\}$ - на рис. 4.3.3 (см. табл. П.7.1). Сопоставление рисунков показывает, что наиболее устойчивой к введению шума оказалась периодограмма Шустера, в то время как совместное действие субгармоник и шума на $\Theta_N(\nu)$ и $L_N(\nu)$ скрыло основное колебание.

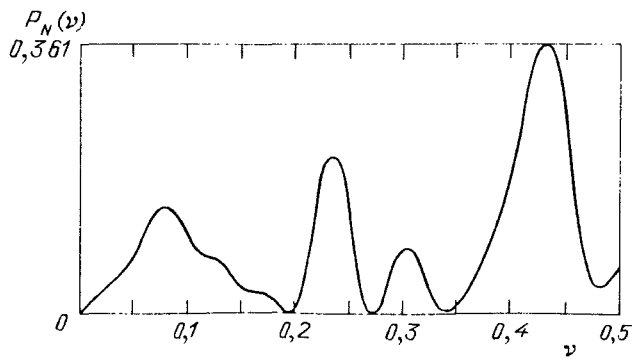


Рис. 4.3.2. Периодограмма отсчетов шума $\{\xi_k\}$

Заметим, что при расчетах в данном примере намеренно бы-

ло выбрано слишком высокое спектральное разрешение ($\delta\nu$ = =0,001), чтобы продемонстрировать реальное разрешение непараметрических методов, обусловленное сохранением взаимного расположения точек на фазовой диаграмме при небольшом изменении пробного периода.

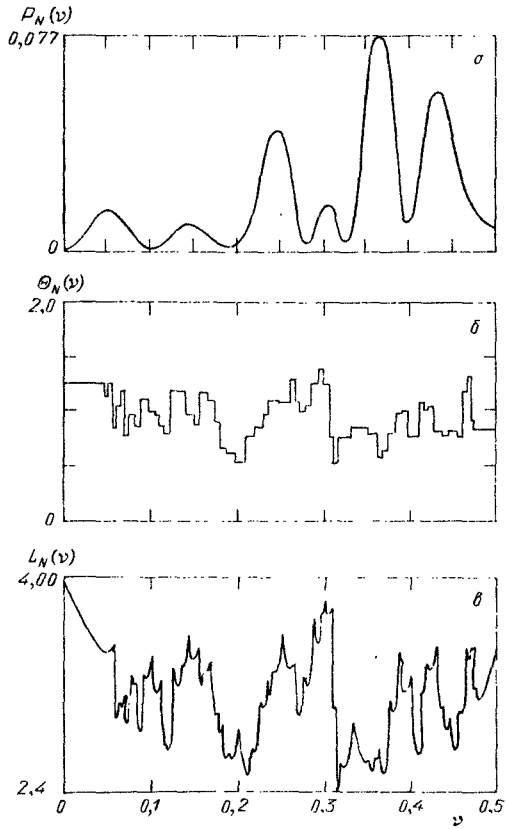


Рис. 4.3.3. Те же периодограммы, что и на рис. 4.3.1, для зашумленных данных ($y_k = f_k + \xi_k$)

Статистика Рексона (1978) представляет собой взвешенную сумму квадратов последовательных разностей ординат точек на фазовой диаграмме:

$$\theta_N^{(R)} = \sum_{k=0}^{N-1} w_k (\vec{Y}_{k+1} - \vec{Y}_k)^2 , \qquad (4.3.5)$$

где коэффициенты

$$w_k = \frac{1}{(x_{k+1} - x_k)^2 + \varepsilon^2} \qquad (4.3.6)$$

выбраны так, чтобы разности ординат соседних точек с большим фазовым расстоянием входили в $\Theta_N^{(R)}$ с малым весом. Этот подход согласуется с теми соображениями, которые были высказаны в начале данного параграфа. В (5) предполагается выполненным условие цикличности (2). Введение постоянной ε преследует цель избежать неоправданно большого веса точек с близкими фазами. При $\varepsilon = 0$ параметр Ренсона представляет собой сумму квадратов коэффициентов наклона всех последовательных отрезков между точками на фазовой диаграмме.

Примеры использования данного критерия содержатся в оригинальной статье Ренсона (1978) и недавней работе Хека и др. (1985), результаты которой обсуждаются в § 4.6.

Метод Пельта (1975, 1980) не предназначен специально для анализа малого количества наблюдений, но может быть использован и в этих условиях. Пельт исходит из общей квадратичной формы

$$\Theta_N^{(P)}(\nu) = \frac{N}{2} \frac{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} g(t_i, t_j, \nu)[y(t_i) - y(t_j)]^2}{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} g(t_i, t_j, \nu)} \qquad (4.3.7)$$

как меры качества представления временного ряда $\{y(t_i)\}$ периодическим процессом с частотой ν. Весовая функция $g(t_i, t_j, \nu)$ не конкретизируется раз и навсегда, она выбирается в зависимости от предварительных сведений о процессе и характера поставленной задачи. Это позволяет, например, использовать статистику Пельта и при быстром анализе данных с целью получить первые указания на возможную периодичность, и при уточнении параметров для уже выделенных периодических колебаний. Требуется, конечно, чтобы весовая функция была не мала лишь тогда, когда фазы моментов t_i и t_j близки. Это условие, естественное для всякого метода выделения периодичности, сближает (7) со статистиками типа ALK, но в методе Пельта ненулевые веса g_{ij} задаются для всех $N(N-1)/2$ различных пар точек, тогда как в методе ALK веса равны 1

для соседних точек на фазовой диаграмме и 0 - для остальных, а в методе Ренсона веса, тоже отличные от нуля только для соседних точек, зависят еще от разностей фаз.

Поскольку весовая функция в методе Пельта выбирается не на основании какого-либо общего принципа, а по желанию исследователя, только значительный опыт практической работы может показать полезность и особенности конкретного представления функции $g(t_i, t_j, \nu)$.

В качестве такого представления Пельт (1980) полагает, в частности,

$$g(t_i, t_j, \nu) = w[\nu(t_i - t_j)] , \qquad (4.3.8)$$

где функция $w(x)$ - четная, периодическая (с периодом 1) и существенно отличается от нуля лишь в окрестности целых значений x, а затем выбирает коэффициенты w_n в разложении Фурье

$$w(x) = \sum_{n=0}^{\infty} \cos(2\pi n x) w_n \qquad (4.3.9)$$

так, чтобы вычисление статистики (7) было эквивалентно полигармоническому отбеливанию в случае равномерной временной последовательности. Можно использовать и представление $w(x)$ в виде ступенчатой функции (Пельт, 1975; Харвиг и Кальв, 1975). По существу такое представление используется всегда, изменяется лишь ширина области ε, в которой $w(x)$ не мала. При значительном ε статистика Пельта близка к обратной периодограмме Шустера, а при $\varepsilon \ll 1$ - к статистике ALK. Пельт замечает также, что иногда в (7) целесообразно использовать вместо квадрата разности ординат модуль этой разности.

Дальнейшие сведения, касающиеся особенностей обсуждаемого метода и способов быстрого вычисления периодограммы $\theta_N^{(P)}$, а также примеры анализа рядов наблюдений переменных звезд и программы расчетов содержатся в книге Пельта (1980).

§ 4.4. Автокорреляционные методы

Для неравномерных временных рядов автокорреляционный метод нельзя применить в том виде, который использовался в

гл. 2 при изучении равномерных рядов. Вместе с тем, если несколько изменить определение отклика на сдвиг ряда вдоль оси времени, то можно найти близкие аналоги автокорреляционной функции. Ниже описаны два варианта такого подхода, которые были предложены в астрономической литературе для исследования колебаний, сохраняющих фиксированное значение периода лишь в течение нескольких циклов. По мнению авторов в случаях таких квазипериодических колебаний автокорреляционные методы чувствительнее стандартных методов поиска периода, использующих сводную информацию о временном процессе. С другой стороны, при наличии сложных колебаний, скажем, характеризующихся двумя периодами, анализировать автокорреляционные кривые значительно труднее, чем распределение спектральной плотности. Этот вопрос нуждается в дальнейшем изучении.

Метод Г.Бэрка, Мэдера и Рафинера (1978). Пусть $\{t_k, y_k\}$, $k = 0, 1, \ldots, N-1$ - произвольный ряд N измерений. Образуем второй временной ряд из N отсчетов $\{t'_k, y'_k\}$ путем сдвига исходного ряда вдоль временной оси на пробный интервал τ :

$$t'_k = t_k + \tau, \quad y'_k = y_k. \tag{4.4.1}$$

Новый ряд сосредоточен на промежутке времени $[t_0 + \tau, t_{N-1} + \tau]$. Рассмотрим теперь выборку всех $2N$ отсчетов, расположенных в порядке возрастания времени, и в частности ту ее часть, которая является общей для исходного и сдвинутого рядов. Общая часть рядов расположена на промежутке $[t_0 + \tau, t_{N-1}]$, она включает некоторое количество $n(\tau) < 2N$ отсчетов y''_j. Наконец, образуем сумму типа (4.1.6):

$$\zeta(\tau) = \frac{\sum_{j=1}^{n(\tau)-1} (y''_{j+1} - y''_j)^2}{2\sum_{j=1}^{n(\tau)} (y''_j - \bar{y}'')^2}. \tag{4.4.2}$$

Если пробный интервал τ мало отличается от квазипериода колебаний P_0, то $\zeta(\tau) \ll 1$; в противном случае значения $\zeta(\tau)$ близки к 1 (при периодическом процессе с высоким отношением сигнала к шуму и большим числом наблюдений за период мы получим $\zeta \ll 1$ и для $\tau \to 0$). Очевидно, менее глубокие минимумы

$\zeta(\tau)$ будут наблюдаться также при $\tau = 2P_0,\ 3P_0, \ldots$

С применениями описанного метода к искусственному (синусоида плюс шум) и действительным рядам наблюдений можно ознакомиться в цитированной выше работе (авторы, впрочем, не учитывают того обстоятельства, что смещение максимума кривой $\zeta(\tau)$ от $P_0/2$ может быть вызвано асимметрией кривой блеска). Сопоставление автокорреляционного метода с другими методами поиска периода производится в работе Хека и др. (1985), результаты которой обсуждаются в § 4.6.

Метод Перси, Джакате и Мэттьюза (1981) заключается в предварительном образовании всевозможных разностей моментов времени и отсчетов временного ряда:

$$\Delta t_{ij} = t_j - t_i\ , \qquad \Delta y_{ij} = y_j - y_i\ , \qquad (4.4.3)$$
$$i = 0,\ 1, \ldots,\ N-2; \quad i < j \leqslant N-1\ .$$

Затем весь интервал изменения Δt_{ij} разбивается на бины подходящей длины вблизи подозреваемого значения периода, и модули разностей $|\Delta y_{ij}|$ усредняются в каждом бине. Если процесс $y(t)$ содержит периодический компонент с основным периодом (или квазипериодом) P_0, то следует ожидать, что зависимость $\overline{|\Delta y(\Delta t)|}$ будет иметь минимумы при значениях Δt, близких к $P_0,\ 2P_0, \ldots$ Заметим, что при необходимости метод легко реализуется вручную путем совмещения начала системы координат $(\Delta t,\ \Delta y)$, нарисованной на прозрачной бумаге, с отсчетами ряда и перенесения на прозрачную бумагу всех последующих отсчетов.

§ 4.5. Сопряженные периоды

Как в анализе Фурье, так и в непараметрических методах выявление истинного периода колебания в значительной мере осложняется наличием дополнительных экстремумов периодограммы, называемых сопряженными или ложными. Появление таких экстремумов отражает общее явление *подмены частот* (aliasing), обусловленное дискретностью совокупности моментов наблюдения непрерывного процесса. Дополнительные трудности связаны с группированием этих моментов с несколькими периодами, чаще всего кратными суткам, месяцу и году.

В астрономическом аспекте явление подмены частот, по-видимому, впервые обсуждалось Бэйли (1901), а вслед за ним многими авторами, из которых мы отметим Хоффмейстера (1914), Шепли (1914) и Теннера (1948). Основное соотношение, связывающее значения сопряженных периодов, было найдено Хагеном (1921).

Мы неоднократно сталкивались с различными сторонами обсуждаемого явления в предшествующем изложении. Так, упомянутая формула Хагена была приведена в § 1.1. Именно, было показано, что совокупность частот, сопряженных данной частоте ν, для периодограммы Шустера равномерного временного ряда имеет вид:

$$\nu_l = \pm\nu \pm l/\Delta t\,, \qquad l = 0,\ 1,\ 2, \ldots, \qquad (4.5.1)$$

где Δt - шаг дискретизации (здесь следовало бы ввести дополнительные индексы, указывающие комбинацию знаков, но мы не будем усложнять обозначения, так как при обобщении (1) надобность в индексах исчезнет). С особой наглядностью явление подмены частот выступает при отсутствии шума (рис. 1.1.7). На периодограмме неравномерного ряда также имеются ложные максимумы, тем более явственные, чем дальше отстоит распределение моментов наблюдения от хаотического (примеры 1.6.2 и 3.2.2). Способы выделения истинной периодичности обсуждались в §§ 3.3 и 3.4. Наконец, общее рассмотрение проблемы подмены частот, основанное на выяснении степени перекрытия спектров при дискретизации исходного процесса, дано в §§ 2.1 и 2.2.

Для наглядного описания совокупности (1) ось частот иногда представляют в виде «гармошки» (рис. 4.5.1) с шагом, равным критической частоте Найквиста $\nu_c = (2\Delta t)^{-1}$. При этом сопряженные частоты отвечают точкам пересечения с осью частот вертикальной прямой, проведенной через заданную точку ν.

Как неоднократно отмечалось выше, периодограмма Шустера характеризует степень соответствия распределения отсчетов на фазовой диаграмме заданному наперед колебанию - синусоиде, оптимальным образом сдвинутой по фазе. Вполне естественно, что для чисто гармонического колебания с периодом

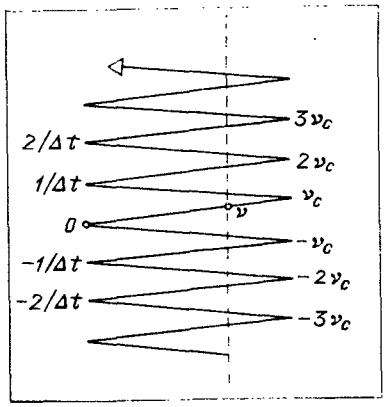


Рис. 4.5.1. Схематическое представление совокупности сопряженных частот

$P_0 \equiv \nu_0^{-1}$, когда спектр содержит лишь две линии на частотах $\pm\nu_0$, сопряженные частоты следуют соотношению (1) при $\nu = \nu_0$. Если же форма колебания отличается от синусоидальной, то его спектр содержит гармоники с частотами $2\nu_0$, $3\nu_0, \ldots$, каждой из которых соответствуют свои сопряженные частоты. Поэтому колебанию с периодом P_0 произвольной формы отвечают на периодограмме Шустера сопряженные частоты вида

$$\nu_{n,l} = \pm n\nu_0 \pm l/\Delta t, \qquad \begin{matrix} n = 1, 2, 3, \ldots, \\ l = 0, 1, 2, \ldots \end{matrix} \qquad (4.5.2)$$

Допуская здесь возможность отрицательных значений индексов, перепишем (2) в более простом виде:

$$\nu_{n,l} = n\nu_0 + l/\Delta t, \qquad \begin{matrix} n = \pm 1, \pm 2, \ldots, \\ l = 0, \pm 1, \pm 2, \ldots, \end{matrix} \qquad (4.5.3)$$

так что совокупность сопряженных периодов определяется соотношением

$$P_{n,l} = \left(\frac{n}{P_0} + \frac{l}{\Delta t} \right)^{-1}. \qquad (4.5.4)$$

Хорошей иллюстрацией к правилу (3) является рис. 1.1.7. На практике основную трудность составляет обычно выбор в качестве истинного одного из колебаний с частотами $\nu_{1,0}$ и $\nu_{-1,1}$ (рис. 4.5.2).

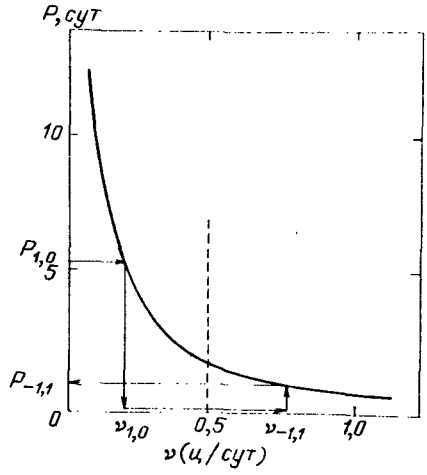


Рис. 4.5.2 Основные сопряженные периоды и частоты

Сворачивание ряда с интервалами времени, кратными действительному периоду P_0, дает фазовую диаграмму, включающую соответственно 2, 3,... полных цикла колебания. Вообще говоря, получающееся при этом распределение точек на диаграмме плохо согласуется с «референтным» колебанием в виде синусоиды, так что периодограмма Шустера не должна обнаруживать максимумы на частотах субгармоник $\nu_0/2$, $\nu_0/3$,... Однако методы поиска периодичности, не связанные наперед заданной формой колебания и учитывающие только внутренний разброс отсчетов около произвольной средней «кривой блеска», обнаружат отклик и на частотах субгармоник. Поэтому наборы сопряженных частот и периодов для непараметрических методов имеют вид

$$\nu_{n,l,r} = \frac{1}{r}\left(n\nu_0 + \frac{l}{\Delta t}\right), \qquad P_{n,l,r} = r\left(\frac{n}{P_0} + \frac{l}{\Delta t}\right)^{-1}, \tag{4.5.5}$$

$$n = \pm 1, \pm 2, \ldots, \quad l = 0, +1, \ldots, \quad r = 1, 2, \ldots$$

Поскольку с ростом r на отрезке, равном общей продолжительности временного ряда $T \equiv t_{N-1} - t_0$, укладывается все меньшее количество отрезков длины rP_0, экстремумы периодограммы с увеличением r убывают по величине.

В качестве примера проявления сопряженных частот на периодограммах Шустера и Аббе-Лафлера-Кинмана укажем на рис. 1.6.1 и 1.7.1, соответствующие ряду оценок блеска

звезды CF Cas. При первом взгляде на рис. 1.7.1 кажется, что наличие такой обширной совокупности сопряженных частот, как это дает (5), ставит непреодолимые препятствия при выделении колебаний с истинным периодом. Однако сопряженные частоты подчиняются определенным сериальным закономерностям, и это обстоятельство весьма полезно при «дешифровке» периодограмм. С другой стороны, непараметрические методы дают на основной частоте ν_0 больший отклик, чем метод Фурье-Шустера (мощность не распределяется по гармоникам). Поэтому при анализе данных наблюдений разумно вычислить периодограммы и того, и другого типа.

Сказанное выше еще раз подчеркивает общее правило: не следует противопоставлять друг другу различные методы выделения периодичности, поскольку они позволяют получить взаимно дополнительную информацию об изучаемом процессе.

Строго говоря, наборы сопряженных частот (3) - (5) отвечают лишь равномерному временному ряду с фиксированным шагом дискретизации Δt. Такие ряды встречаются в астрономической практике, скажем, при измерении колебаний блеска переменной звезды с периодом порядка 1 часа в течение ночи. Однако чаще всего астрономические ряды лишь приближенно равномерны; они, можно сказать, *модулированы* некоторой частотой $\tilde{\nu}$, значение которой для конкретного ряда может отличаться от обратных суток, месяца и года. Предположим, например, что в течение 20 последовательных ночей производились измерения блеска периодической звезды вблизи меридиана. Значения моментов наблюдений t_0, t_1, ..., t_{19} не образуют строго равномерного ряда, и поэтому наилучшим равномерным приближением к этой последовательности будет ряд с шагом, несколько отличным от звездных суток. С увеличением количества измерений следует ожидать, что периоды модуляции будут все ближе к их теоретическим значениям.

Таким образом, на практике следует пользоваться формулами для сопряженных частот

$$\nu_{n,l} = n\nu_0 + l\tilde{\nu} \qquad (4.5.6)$$

и

$$\nu_{n,l,r} = \frac{1}{r}\,(n\nu_0 + l\tilde{\nu}), \qquad \begin{matrix} n = \pm 1,\ \pm 2,\ \ldots\ ; \\ l = 0,\ \pm 1,\ \ldots\ ; \\ r = 1,\ 2,\ \ldots \end{matrix} \qquad (4.5.7)$$

при употреблении соответственно параметрических и непараметрических методов, причем значение частоты модуляции $\tilde{\nu}$ подбирается наилучшим образом в каждом конкретном случае. Это легко сделать, измерив расстояние между экстремумами периодограммы, отвечающими сопряженным частотам.

Если наблюдается объект, расположенный вне Солнечной системы, то частота модуляции $\tilde{\nu}$ будет близка к $1/D_*$, где D_* - продолжительность *звездных суток*. С другой стороны, время обычно измеряется в *солнечных сутках* $D_\odot$, так что при измерении частот в единицах цикл/солнечные сутки (цикл/с.сут) следует принять $\tilde{\nu} = D_\odot/D_* = 1{,}00274$ цикл/с.сут (Теннер, 1948). Аналогично для частоты модуляции, отвечающей *синодическому лунному месяцу* (от новолуния до новолуния), получаем значение $\tilde{\nu} = 1/29{,}530\ 588$ с.сут $= 0{,}033\ 863$ цикл/с.сут, а для частот, отвечающих *тропическому* (от равноденствия до равноденствия) и *сидерическому* (относительно неподвижных звезд) *годам* получаем соответственно $\tilde{\nu} = 1/365{,}2422$ с.сут $= 0{,}002\ 737\ 91$ цикл/с.сут и $\tilde{\nu} = 1/365{,}2564$ с.сут $= 0{,}002\ 737\ 80$ цикл/с.сут.

§ 4.6. Сопоставление различных методов поиска скрытой периодичности

Практика показывает, что при выборе метода анализа данных исследователи чаще руководствуются собственными склонностями, чем характером изучаемой задачи. Между тем, как нетрудно видеть, эффективность какого-либо метода существенным образом зависит от структуры процесса и условий его наблюдения.

Предположим, например, что задача состоит в отыскании периодического компонента в ряду измерений блеска переменной звезды. Классический анализ Фурье, предназначенный, вообще говоря, для решения общей задачи оценки спектра мощности, может быть с успехом использован и при изучении лишь дискретной части спектра, обусловленной периодическими ком-

понентами. Однако вполне вероятна ситуация, когда анализ Фурье не выявит периодического компонента, в то время как это оказывается возможным при помощи другого метода. Пусть, скажем, периодический компонент имеет сложный, далекий от синусоидального вид, подобный кривым блеска затменных переменных звезд. Тогда его спектр будет включать целый ряд дискретных линий, причем не исключено, что наиболее интенсивными будут далекие члены этого ряда. Ясно, что присутствие даже незначительного шума может полностью скрыть дискретную составляющую спектра. В этих условиях более эффективным окажется применение одного из непараметрических методов, суммирующих на частоте периодичности информацию о всем дискретном спектре мощности.

Приведенный пример хорошо иллюстрирует и такую возможность, когда эффективным окажется «сверхпараметрический» метод, сравнивающий изучаемый ряд с некоторым наперед заданным несинусоидальным колебанием. Действительно, в качестве отклика на пробное колебание мы могли бы рассматривать уменьшение суммы квадратов уклонений членов ряда от функции типа кривой блеска затменных переменных звезд, содержащей лишь несколько подгоночных параметров (см. § 1.7). При определенных значениях этих параметров отклик будет существенно превосходить максимальное значение периодограммы Шустера, создавая иллюзию чрезвычайно высокой чувствительности метода. Однако последняя обусловлена лишь наличием *дополнительной информации* об изучаемом процессе, которая может быть надлежащим образом включена и в анализ Фурье. Как отмечалось в § 1.7, гибкость метода Фурье связана с тем обстоятельством, что тригонометрические функции образуют *полную* систему, по которой можно разложить колебание произвольной формы.

В геофизике и радиофизике за последнее десятилетие накоплен значительный опыт работы с методом максимума энтропии и родственными ему методами (§§ 2.6, 2.5). Информационный подход к задаче оценивания спектра мощности оказался весьма плодотворным, и в тех случаях, когда его можно было применить, показал свое превосходство над традиционным подходом. К сожалению, область применимости метода максимума

энтропии ограничена равномерными временными рядами достаточно редкими в астрономической практике.

Таким образом, при сопоставлении различных методов поиска периодичности следует говорить не об отвлеченной их эффективности, а об относительных преимуществах и недостатках в условиях, типичных для астрономических наблюдений. Основными факторами, определяющими выбор метода исследования, являются:

а) отношение сигнала к шуму;

б) общее количество отсчетов N;

в) среднее число измерений, приходящихся на один период:

$$\mathcal{H} = \frac{NP_0}{T}, \qquad (4.6.1)$$

где T - длина ряда;

г) форма периодического колебания;

д) детальный характер временного распределения отсчетов (спектр скважности);

е) наличие априорных сведений о значении периода и форме колебания.

В настоящем параграфе мы приведем результаты двух специальных исследований рассматриваемого вопроса - Хека и др. (1985) и Пельта (1980). В первом из них сначала аналитически формулируется основная задача при оценивании уровня значимости статистик в методах Фурье (F), автокорреляционном Бэрка и др. (A), Ренсона (R), Стеллингвефа (S) и Аббе-Лафлера-Кинмана (ALK); затем перечисленные методы применяются к искусственному временному ряду, включающему периодический компонент. Во втором автор суммирует опыт анализа реальных и модельных временных рядов с точки зрения исследователя, интересующегося лишь поиском периодических составляющих.

Хек и др. (1985) задают модель следующим образом. Протяженность временного ряда составляет 20 суток, из которых исключены две-три ночи (имитация плохих погодных условий). Рассматривались колебания с периодами P = 0,75; 0,95; 1,92; 1,98; 3,5 суток; второе из этих значений специально принято близким к периоду модуляции - 1 суткам, а третье и четвер-

тое - к 2 суткам (частота периодичности лежит вблизи предела Найквиста, см. рис. 4.5.2). Для полного числа измерений выбраны три значения: N = 15, 28 и 50. Были предприняты специальные меры с целью приблизить распределение моментов измерения к условиям реальных астрономических наблюдений. Кривая блеска задавалась полиномом 3-й степени так, чтобы получились нормированные к единичной амплитуде колебания разного вида: симметричное (0), асимметричное (1) и типа двойной волны (2). Наконец, в модель вводился аддитивный гауссовский шум с тремя значениями стандартного отклонения σ = 0,1, 0,3 и 0,5. Для каждого набора перечисленных параметров были реализованы 3 модельных ряда, каждый из которых последовательно анализировался при помощи 5 перечисленных выше методов. Эффективность методов оценивалась по степени выраженности экстремума соответствующей периодограммы.

Для небольшого количества измерений лучшие результаты дает метод R и худшие - A, при большом числе отсчетов метод S слегка превосходит методы A и F. По мере увеличения уровня шума σ первенство переходит от R к S; определенно можно сказать, что метод A не предназначен для работы с низким отношением сигнала к шуму. Для кривых блеска типов 0 и 1 лучшие результаты дает метод F и худшие - A; для кривых типа 2 систематически лучше R, в то время как F часто обнаруживает полупериод двойной волны. Что касается точности выделения периодов, то здесь методы R и S превосходят F и A (за исключением случая P = 1,98 сут, когда F превосходит все остальные методы - возможно, из-за объединения истинного и сопряженного пиков в один широкий максимум при P = 2 сут). В отношении скорости счета все методы кроме S приблизительно эквивалентны; относительно большая продолжительность вычислений в методе S обусловлена способом группирования данных. Малая эффективность автокорреляционного метода вполне понятна, поскольку он был разработан специально для изучения квазипериодических колебаний.

В общем, при заданных условиях методы R и S показывают несколько более высокую эффективность по сравнению с методами F и A, однако, по мнению авторов, вывод о безусловном превосходстве непараметрических методов поиска периодичнос-

ти был бы слишком поспешным.

Пельт (1980) сопоставляет методы Аббе-Лафлера-Кинмана, Ренсона, Юркевича, LS_m-спектр, полигармоническое отбеливание и различные варианты собственной статистики (4.3.7). Он полагает, что на первоначальном этапе поиска периодичности следует использовать формулы (4.3.7) и (4.3.9) при нескольких отличных от нуля первых коэффициентах w_n. Ступенчатое представление функции $w(x)$ в (4.3.8) позволяет получить универсальный алгоритм поиска и уточнения периода. Остальные перечисленные выше методы целесообразно применять для анализа сравнительно небольшого количества измерений либо для нахождения более точного значения периода.

Несомненно, вопрос о сравнительной эффективности методов выявления периодических колебаний нуждается в дальнейшем изучении. Следует лишь еще раз подчеркнуть, что современные компьютеры позволяют без труда реализовать несколько алгоритмов для исследования одних и тех же данных, и потому в практической работе неразумно противопоставлять различные подходы.

ГЛАВА 5

ИМПУЛЬСНЫЕ СЛУЧАЙНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ

Типичная задача, обсуждавшаяся в предыдущих главах, формулируется следующим образом. Некоторый непрерывный процесс $y(t)$ наблюдается в моменты времени $t_0, t_1, \ldots, t_{N-1}$, не обязательно равноудаленные друг от друга. Исследователю известен только временной ряд $\{t_k, y(t_k)\}$, и требуется как можно более полно восстановить свойства исходного процесса $y(t)$.

Такая постановка проблемы встречается часто, но она далеко не всегда достаточна для описания ситуаций, с которыми приходится встречаться в современных исследованиях. В качестве примера укажем на ряды, получающиеся в экспериментах по счету отдельных частиц - фотонов, нейтрино, а также импульсов механического напряжения в детекторах гравитационных волн. Характерная особенность этих экспериментов заключается в том, что детектор большую часть времени находится в невозбужденном состоянии, и лишь иногда выходит из него на короткий промежуток времени, в течение которого регистрируется импульс. Различие амплитуд импульсов зачастую обусловлено второстепенными или мешающими обстоятельствами, так что простая стандартизация импульсов приводит к последовательности идентичных кратковременных событий, случившихся в моменты $\{t_k\}$.

Мы пришли к понятию *точечного случайного процесса*, применимому в тех случаях, когда протяженность отдельных событий пренебрежимо мала по сравнению со средним интервалом времени между событиями. Точечный процесс задан лишь случайной последовательностью моментов регистрации событий $\{t_k\}$.

Основная область астрофизических приложений точечных процессов - внеатмосферные исследования в рентгеновском диапазоне. В течение последней четверти века счет фотонов благодаря своим принципиальным преимуществам перед другими способами регистрации стал основным методом наблюдений и в оптической области спектра. Очевидно, если изучаются колебания светимости объекта с характерным временем τ, значительно превосходящим среднее время между регистрацией фотонов $\overline{X}$, то количество событий, случившихся за время экспозиции, можно рассматривать как непрерывную величину, и мы возвращаемся к обычному описанию процесса. В области же $\tau \sim \overline{X}$ на передний план выступают особенности, присущие импульсным процессам.

Следует заметить, что при рассмотрении точечных последовательностей скважность наблюдений обычно мешает сравнительно мало, т.е. почти всегда имеются интервалы непрерывного слежения, содержащие много событий. Поэтому основную трудность составляет не учет дискретности временного ряда, а извлечение информации из *случайной по самой природе последовательности событий.*

Значительная часть излагаемого в данной главе аппарата описана в математической и физической литературе. Так сложилось, что некоторые методы были переоткрыты в астрономии, а другие остаются малоизвестными. К последним относится и важное понятие импульсного спектра $I_N(\nu)$, который пропорционален оценке спектральной плотности средней интенсивности процесса (из общих соображений следует ожидать, что в качестве такой оценки выступит периодограмма Шустера $P_N(\nu)$ ряда $\{t_k, 1\}$, полученного приписыванием равных весов всем событиям). Общие методы, предназначенные для анализа произвольных точечных процессов, излагаются в §§ 5.2 - 5.4; в последующих параграфах рассматриваются характерные вопросы, возникающие при реализации систем счета событий.

§ 5.1 Пуассоновский процесс

Исключительная роль, которую играет *процесс Пуассона* в научных и технических исследованиях, обусловлена несколькими причинами. Прежде всего, необходимые для формирования

пуассоновского процесса условия столь естественны, что они очень часто выполняются с хорошим приближением. Так, при помощи пуассоновского процесса удается описать последовательности вызовов на телефонной станции, распада радиоактивных ядер, регистрации фотонов приемником и т.д. Кроме того, в определенном смысле пуассоновский процесс можно назвать «совершенно случайным», так что он является референтным процессом для проверки выводов относительно реальности группирования событий (вспышек), периодичности и прочих явлений детерминированного характера в реальных импульсных процессах.

Свойства пуассоновского процесса описаны во многих математических и прикладных руководствах, в частности, Гнеденко (1988), Феллера (1966), Кокса и Льюиса (1966), Рытова (1976), А.В. и А.А.Солодовых (1988). Поэтому мы кратко остановимся на характеристиках классического процесса с постоянной средней интенсивностью событий, а затем рассмотрим случаи, когда средняя интенсивность детерминированно или стохастически изменяется со временем.

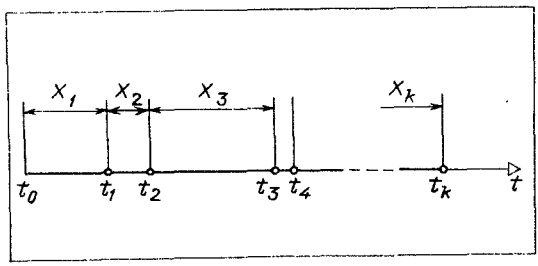


Рис. 5.1.1. Распределение событий и интервалов времени между ними для точечного случайного процесса

Однородный пуассоновский процесс. Предположим, что на бесконечной временной оси (рис. 5.1.1) отмечены моменты наступления некоторых событий, протяженность которых пренебрежимо мала по сравнению со средним интервалом времени между последовательными событиями (т.е. события считаются «точечными»). Совокупность $\{t_k\}$ моментов наступления событий предполагается случайной (ср. для определенности вызовы на коммутаторе), поэтому число событий $\mathfrak{N}(t_0, t)$, происшедших

на промежутке $[t_0, t]$, есть случайная величина, которая может принимать дискретные значения 0, 1,. ., k,... с некоторыми пока неопределенными вероятностями

$$\Pr[\mathfrak{N}(t_0, t) = k] = p_k(t_0, t) \ . \qquad (5.1.1)$$

Найдем вероятностное распределение $\{p_k\}$ при следующих трех условиях. I). События ординарны, т.е. в течение малого промежутка времени Δt вероятность наступления двух или более событий пренебрежимо мала:

$$p_0(t_0, t_0 + \Delta t) \simeq 1 - n\Delta t,$$

$$p_1(t_0, t_0 + \Delta t) \simeq n\Delta t, \qquad (5.1.2)$$

$$p_k(t_0, t_0 + \Delta t)/\Delta t \to 0, \quad k \geqslant 2, \ \Delta t \to 0,$$

где n - параметр с размерностью, обратной времени. II). Вероятности (1) не зависят от конкретного выбора начального момента t_0, а определяются лишь продолжительностью интервала наблюдений $t - t_0$. Таким образом, процесс однороден по времени (его статистические свойства не изменяются при временном сдвиге), и в частности, параметр n - постоянная величина. Полагая для простоты $t_0 = 0$ и обозначая $\mathfrak{N}(0, t)$ через $\mathfrak{N}(t)$, имеем вместо (1):

$$\Pr[\mathfrak{N}(t) = k] = p_k(t), \qquad k = 0, 1, \ldots \qquad (5.1.3)$$

III). Вероятность наступления события в промежутке $[t, t+\Delta t]$ не зависит от того, что произошло до момента t.

Эти три условия - *ординарность*, *однородность* и *отсутствие последействия* - определяют *однородный пуассоновский процесс* и позволяют найти $p_k(t)$ в явном виде. Действительно, вероятность наступления k событий в интервале $[0, t+\Delta t]$ можно записать на основании I-III следующим образом:

$$p_0(t + \Delta t) \simeq p_0(t)p_0(\Delta t) \simeq p_0(t)(1 - n\Delta t),$$

$$p_k(t + \Delta t) \simeq p_k(t)p_0(\Delta t) + p_{k-1}(t)p_1(\Delta t) \simeq \qquad (5.1.4)$$

$$\simeq p_k(t)(1 - n\Delta t) + p_{k-1}(t)n\Delta t, \quad k = 1, 2, \ldots$$

Устремляя здесь Δt к нулю, мы приходим к простым дифференциальным уравнениям

$$\frac{dp_0(t)}{dt} = -np_0(t),$$

$$\frac{dp_k(t)}{dt} = -np_k(t) + np_{k-1}(t), \quad k = 1, 2, \ldots, \tag{5.1.5}$$

которые последовательно решаются при заданных начальных условиях $p_0(0) = 1$, $p_k(0) = 0$, $k = 1, 2, \ldots$ Как легко видеть,

$$p_k(t) = e^{-nt}\frac{(nt)^k}{k!}, \quad k = 0, 1, \ldots \tag{5.1.6}$$

Итак, для однородного пуассоновского процесса вероятность наступления k событий в течение интервала длительности t определяется выражением (6). Поскольку случайная величина $\mathfrak{N}(t)$ зависит от параметра t, мы можем трактовать ее в соответствии с § 1.5 как случайную функцию времени или случайный процесс. Две реализации этого процесса приведены на рис. 5.1.2. Функция $\mathfrak{N}(t)$, очевидно, сохраняет свое значение между моментами $\{t_k\}$ наступления событий и испытывает положительные единичные скачки в моменты $\{t_k\}$.

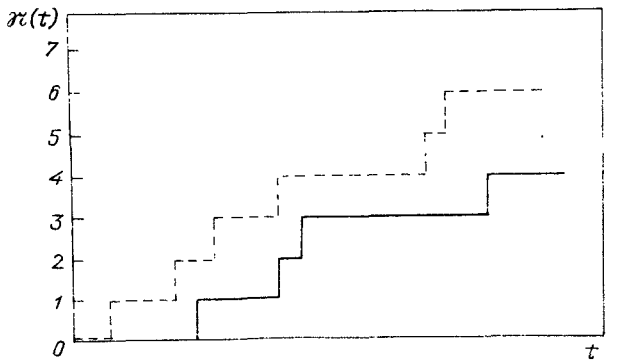


Рис. 5.1.2. Случайные реализации однородного пуассоновского процесса $\mathfrak{N}(t)$

Другой подход, приводящий к распределению Пуассона, рассматривает (6) как предельную форму *биномиального распределения* $C_M^k w^k (1 - w)^{M-k}$ в случае, когда полное число попыток M реализовать некоторое событие в течение интервала длительности t стремится к бесконечности, вероятность w успешной

реализации одного события стремится к нулю, а их произведение Mw - ожидаемое число успешных реализаций в $[0,t]$ - остается постоянным и равным nt (в этой связи формулу Пуассона называют иногда *законом редких событий*).

Производящая функция вероятностей распределения Пуассона находится путем подстановки (6) в формулу (П.4.3):

$$G(s,t) \equiv \sum_{k=1}^{\infty} p_k(t) s^k = e^{-nt(1-s)}, \quad -1 \leqslant s \leqslant 1. \tag{5.1.7}$$

Распределение (6), как и всякое вероятностное распределение, нормировано к единице:

$$\sum_{k=1}^{\infty} p_k(t) = G(1,t) = 1. \tag{5.1.8}$$

Введем обозначения $M(t)$ и $D(t)$ соответственно для среднего числа и дисперсии числа событий, происшедших в $[0,t]$. Согласно (7), (П.4.4) и (П.4.11) имеем:

$$M(t) \equiv \langle \mathfrak{N}(t) \rangle = nt \tag{5.1.9}$$

$$D(t) \equiv \langle \mathfrak{N}^2(t) \rangle - \langle \mathfrak{N}(t) \rangle^2 = nt. \tag{5.1.10}$$

Как и следовало ожидать, среднее число событий линейно растет с увеличением длины интервала наблюдения за процессом. Из (9) следует и простая интерпретация параметра n - это среднее число событий, возникающих в единицу времени, т.е. *средняя интенсивность процесса*.

Две последние формулы выражают характерное свойство пуассоновского процесса: для него дисперсия количества событий равна среднему его значению или, что то же самое, стандартное отклонение равно квадратному корню из ожидаемого числа событий. Именно это широко известное свойство и служит чаще всего на практике для грубой проверки соответствия изучаемого процесса пуассоновскому. Если, например, количество импульсов фотоумножителя $\mathfrak{N}(t)$, порожденных излучением достаточно яркой звезды, флуктуирует около некоторого среднего значения M так, что стандартное отклонение $\sigma_{\mathfrak{N}} \simeq \sqrt{M}$, то наблюдатель вправе считать тракт регистрации близким к линейному счетчику (см. подробнее §§ 5.5, 5.6).

Поскольку однородный пуассоновский процесс служит рефе-

рентным процессом при анализе случайных последовательностей различной природы, полезно ввести для произвольного процесса $\xi(t)$ параметр, равный отношению дисперсии к среднему значению:

$$\nu_{\xi}(t) \equiv \frac{D_{\xi}(t)}{\langle \xi(t) \rangle}. \qquad (5.1.11)$$

Величина ν называется *индексом дисперсии* (Кокс и Льюис, 1966). Согласно (9) и (10). для однородного пуассоновского процесса индекс дисперсии тождественно равен единице:

$$\nu_{\mathfrak{N}}(t) \equiv 1, \qquad (5.1.12)$$

так что отличие $\nu(t)$ от 1 может рассматриваться как мера отклонения изучаемой последовательности от «совершенно случайной» - пуассоновской последовательности событий (см., впрочем, § 5.5).

При $M(t) \gg 1$ пуассоновское распределение (6) приближенно описывается формулой Гаусса (П.4.23) со средним значением и дисперсией, равными $M(t)$.

Характеристики однородного пуассоновского процесса $\mathfrak{N}(t)$ не зависят от того, совпадает начало отсчета времени с некоторым событием или располагается между ними. Предположим, что выполняется первое условие. Тогда вероятность того, что следующее событие произойдет между τ и $\tau+\Delta\tau$, равна

$$p_0(\tau)\ p_1(\Delta\tau) \simeq e^{-n\tau}\ n\Delta\tau. \qquad (5.1.13)$$

Отсюда видно, что непрерывная случайная величина X - длина интервала между последовательными событиями - подчиняется экспоненциальному распределению:

$$\Pr[\tau < X \leqslant \tau+d\tau] \equiv f(\tau)d\tau = e^{-n\tau}\ nd\tau,\ 0 \leqslant \tau < \infty. \qquad (5.1.14)$$

Среднее значение длины свободного интервала

$$\langle X \rangle = \int_0^{\infty} \tau f(\tau)d\tau = \frac{1}{n}, \qquad (5.1.15)$$

ее дисперсия -

$$D_X = \int_0^{\infty} \tau^2 f(\tau)d\tau - \langle X \rangle^2 = \frac{1}{n^2}. \qquad (5.1.16)$$

Интервалы времени между последовательными событиями взаимно независимы.

Мы упоминали выше, что распределение Пуассона часто называют «совершенно случайным». Причина этого заключается в следующем. Предположим, что нас интересует распределение моментов времени наступления событий $\{t_k\}$ при условии, что полное число событий, происшедших в интервале длины t, известно. Можно показать (см., например, Гнеденко, 1988), что при этом моменты появления событий независимы друг от друга и распределены равномерно на интервале $[0,t]$. Поскольку равномерное распределение обладает максимальной дифференциальной энтропией среди всех заданных на конечном интервале распределений (см. Агекян, 1974 и § 2.5), можно сказать, что распределение Пуассона характеризуется наибольшей степенью неопределенности совокупности моментов $\{t_k\}$.

Для сопоставления реальных процессов и пуассоновской последовательности событий нужно уметь моделировать последнюю при различных значениях средней интенсивности n. Это весьма просто сделать, если учесть экспоненциальный вид распределения (14) интервалов между событиями. Действительно, экспоненциальная случайная величина X со средним n^{-1} выражается через равномерно распределенную на промежутке $[0,1]$ случайную величину α формулой (П.4.22), т.е. в данном случае $X = (\ln\alpha)/n$. Поскольку датчики или алгоритмы образования равномерных псевдослучайных чисел есть на всех компьютерах, пуассоновская последовательность формируется реализацией совокупности длин свободных интервалов $\{X_k\}$.

Неоднородный пуассоновский процесс отвечает сохранению свойств I и III - т.е. ординарности событий и отсутствию последействия, но отказу от постоянства средней интенсивности возникновения событий с течением времени. Так, вызовы на телефонной станции, рассматриваемые, скажем, в течение получаса, могут быть приближенно описаны однородным процессом Пуассона, однако интенсивность поступления заказов, оставаясь случайной, явно испытывает колебания среднего значения в течение суток.

Для неоднородного пуассоновского процесса число событий, происшедших в течение интервала $[t_0,t]$, зависит от t_0, и потому мы вынуждены вернуться к общему определению (1), а в соотношениях (2) считать n функцией t_0. Как легко видеть,

при $n = n(t)$ остаются справедливыми и соотношения (4), (5), из которых следует обобщение классической формулы Пуассона (6) на случай неоднородного процесса:

$$p_k(t_0,t) = e^{-M(t_0,t)} \frac{[M(t_0,t)]^k}{k!}, \quad k = 0, 1, \ldots, \qquad (5.1.17)$$

где

$$M(t_0,t) = \langle \mathfrak{N}(t_0,t) \rangle = \int_{t_0}^{t} n(t')dt' \qquad (5.1.18)$$

- среднее число событий, происшедших в интервале $[t_0,t]$. Полагая

$$\bar{n}(t_0,t) = \frac{1}{t - t_0} \int_{t_0}^{t} n(t')dt', \qquad (5.1.19)$$

формуле (17) можно придать вид

$$p_k(t_0,t) = e^{-(t-t_0)\bar{n}} \frac{[(t-t_0)\bar{n}]^k}{k!}, \quad k = 0, 1, \ldots, \qquad (5.1.20)$$

особенно близкий привычной формуле Пуассона (6). Величина (19) соответствует двойному усреднению - статистическому и интервальному.

Продолжим обобщение классического процесса Пуассона. Выражение (7) для производящей функции принимает теперь вид

$$G(s;t_0,t) = e^{-(1-s)M(t_0,t)}, \quad -1 \leqslant s \leqslant 1, \qquad (5.1.21)$$

так что равенство среднего значения $M(t_0,t)$ и дисперсии числа событий $D(t_0,t)$ сохраняется для неоднородного процесса, а индекс дисперсии $\nu_{\mathfrak{N}}(t_0,t)$ по-прежнему равен 1.

Особый интерес представляет распределение длины свободного интервала между событиями. Предположим, что мы начали наблюдение процесса в момент t, когда произошло некоторое событие, и нас интересует X_t - время ожидания до следующего события. Мы имеем

$$\Pr[\tau < X_t \leqslant \tau + d\tau] \equiv f_t(\tau)d\tau = p_0(t,t+\tau) \cdot p_1(t+\tau, t+\tau+d\tau) =$$

$$= e^{-M(t,t+\tau)} n(t+\tau)d\tau. \qquad (5.1.22)$$

Это выражение определяет плотность распределения интервала

между событиями для произвольной зависимости средней интенсивности от времени. Нетрудно понять также, что начало наблюдения могло и не совпадать с каким-либо событием, поэтому (22) дает плотность распределения времени ожидания до наступления события, начиная с произвольно выбранного исходного момента.

Плотность $f_t(\tau)$ характеризует локальное распределение времени ожидания. Если наблюдения продолжались в течение промежутка длительности T, то суммарное распределение длин интервалов между событиями получается усреднением (22) с весом $n(t)$, т.е.

$$\overline{f}_{t_0}(\tau) = \frac{1}{\overline{n}T}\int_{t_0}^{t_0+T} n(t)n(t+\tau)e^{-M(t,t+\tau)}dt, \qquad (5.1.23)$$

где $\overline{n}$ - средняя по интервалу $[t_0, t_0+T]$ интенсивность событий (19). При $n(t) \equiv n = \text{const}$ формула (23) определяет в согласии с (14) экспоненциальное распределение длин свободных интервалов.

Следует заметить, что при достаточно быстром убывании $n(t)$ может случиться так, что за бесконечный промежуток времени ожидаемое число событий $M(t_0,\infty)$ окажется конечным. При этом в (17) возможны лишь значения $k \leqslant M(t_0,\infty)$. Мы не будем рассматривать этот малоинтересный случай, поэтому всюду ниже предполагается $M(t_0,\infty) = \infty$.

Для моделирования неоднородного пуассоновского процесса нужно реализовать последовательность интервалов $\{X_{t_0}^{(k)}\}$ с плотностью распределения (22), что в общем случае сделать не столь просто, как при $n = \text{const}$. Лишь иногда здесь удается воспользоваться непосредственно соотношением (П.4.18). Для произвольной функции $n(t)$ величину X_t можно реализовать при помощи остроумного метода Дж. фон Неймана (см., например, Ермаков и Михайлов, 1982; Пресс и др., 1986), основная идея которого заключается в следующем. Выберем произвольную простую функцию $g(\tau) \geqslant f_t(\tau)$ так, чтобы случайную величину η с плотностью, пропорциональной $g(\tau)$, можно было легко реализовать. Тогда пары точек $[\eta', \alpha' g(\eta')]$, где η' и α' - реализации η и равномерной на $(0,1)$ величины α, равномерно заполнят двумерную область, расположенную под кривой $g(\tau)$.

Следовательно, они равномерно заполнят и область под кривой $f_t(\tau)$, так что абсциссы тех точек, для которых выполняется неравенство

$$\alpha' g(\eta') \leqslant f_t(\eta'), \tag{5.1.24}$$

и образуют реализации величины с заданной плотностью $f_t(\tau)$.

Пример 5.1.1. Простейшим неоднородным процессом является пуассоновский процесс, средняя интенсивность которого линейно зависит от времени:

$$n(t)= n(t_0)\left(1 \pm \frac{t-t_0}{t_*}\right), \quad t \geqslant t_0, \tag{5.1.25}$$

где в случае убывания $n(t)$ значения t ограничены интервалом $[t_0, t_0+t_*]$. Для этого линейно-неоднородного процесса мы имеем из (18)

$$M(t_0, t) = n(t_0)(t-t_0)\left(1 \pm \frac{t-t_0}{2t_*}\right), \tag{5.1.26}$$

а остальные характеристики следуют из приведенных выше формул.

Пример 5.1.2. В классе периодических неоднородных процессов простейшим является, по-видимому, ступенчатый пуассоновский процесс, для которого средняя интенсивность событий скачками изменяется между значениями n_1 и $n_2 > n_1$ (рис. 5.1.3). Обозначая посредством δ долю периода P_0, в течение которой средняя интенсивность сохраняет максимальное значение, находим из (19), что при $t/P_0 \to \infty$ средняя интенсивность процесса стремится к постоянному значению

$$\bar{n} = (1-\delta)n_1 + \delta n_2. \tag{5.1.27}$$

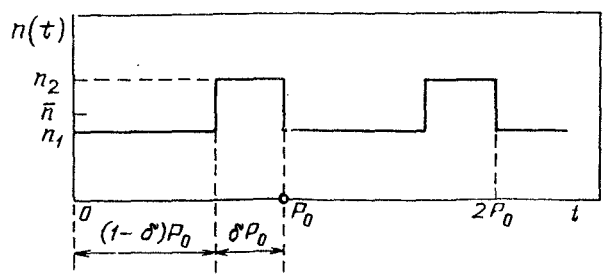


Рис. 5.1.3. Колебания средней интенсивности ступенчатого пуассоновского процесса

При этом вероятностное распределение числа событий на большом интервале времени $t \gg P_0$ стремится к простому закону Пуассона (6) с параметром $\bar{n}$.

Пример 5.1.3. Назовем гармоническим пуассоновским процессом такой неоднородный процесс, для которого средняя интенсивность наступления событий изменяется с течением времени по закону

$$n(t) = \bar{n}\,[(1+\mu\cos(2\pi t/P_0-\varphi_0)], \quad t \geq t_0, \tag{5.1.28}$$

где $\bar{n}$ - усредненная по большому интервалу времени интенсивность, $P_0 \equiv \nu_0^{-1}$, φ_0 и μ - соответственно период, начальная фаза и глубина модуляции колебаний средней интенсивности ($0 \leq \mu \leq 1$). В данном случае

$$M(t_0,t) = \bar{n}\left\{(t-t_0)+\frac{\mu}{2\pi\nu_0}\left[\sin(2\pi\nu_0 t-\varphi_0)-\right.\right.$$
$$\left.\left.-\sin(2\pi\nu_0 t_0-\varphi_0)\right]\right\}. \tag{5.1.29}$$

Важным параметром является $\gamma \equiv \bar{n}P_0 = \bar{n}/\nu_0$ -среднее число событий, происходящих в течение одного периода колебания средней интенсивности.

Гармонический пуассоновский процесс реализуется методом фон Неймана, если принять

$$g(\tau) = \bar{n}(1+\mu)e^{-\bar{n}(1-\mu)\tau} \tag{5.1.30}$$

и получать сначала обычным образом реализации экспоненциальной случайной величины η с плотностью $g(\tau)(1-\mu)/(1+\mu)$, а затем отбирать из них те, для которых выполняется неравенство (24) при законе $f_t(\eta')$, заданном формулами (22), (28) и (29).

Дважды стохастический пуассоновский процесс соответствует средней интенсивности $n(t)$, которая сама является стационарным случайным процессом со средним $\langle n(t)\rangle = \bar{n}$, дисперсией σ_n^2 и автоковариационной функцией

$$\psi_n(\tau) = \langle n(t)n(t+\tau)\rangle - \bar{n}^2, \quad \psi_n(0) = \sigma_n^2. \tag{5.1.31}$$

Именно пуассоновский процесс со случайной интенсивностью описывает последовательность регистраций при счете фотонов в теории излучения (§ 5.5). В этой связи свойства дважды

стохастического процесса в последние годы изучались как аналитически, так и путем численного моделирования (см., например, обзор Меты, 1970). Вообще, как отмечают Кокс и Льюис (1966), «в ситуациях, когда последовательность событий образует приблизительно пуассоновский процесс, подобный способ задания изменения средней интенсивности является, по-видимому, наиболее общим».

При указанных выше предположениях процесс однороден по времени, однако теперь

$$N(t) \equiv \int_0^t n(t')dt' \qquad (5.1.32)$$

представляет собой *случайную* величину, принимающую значения в интервале $[0,\infty)$. Обозначим плотность распределения $N(t)$ через $f_N(x,t)$. Непосредственно из (32) следует, что среднее значение

$$\langle N(t)\rangle = \bar{n}t, \qquad (5.1.33)$$

а дисперсия

$$D_N(t) = 2\int_0^t (t-\tau)\psi_n(\tau)d\tau. \qquad (5.1.34)$$

Найдем распределение числа событий $\{p_k(t)\}$, случившихся в интервале времени длины t. Фиксируем какую-либо реализацию процесса $n(t)$ на указанном интервале; для этой реализации мы имеем неоднородный по времени пуассоновский процесс, распределение вероятностей которого задается выражениями (17) и (18). Теперь примем во внимание стохастический характер $n(t)$, приводящий к тому, что $M(t_0,t)$ в (17) - случайная величина. Ясно, что искомое распределение $\{p_k(t)\}$ получится, если (17) усреднить по всем значениям $M(t_0,t)$. Плотность распределения последней величины, совпадающей фактически с (32), была обозначена выше через $f_N(x,t)$, так что

$$p_k(t) = \int_0^\infty e^{-x}\frac{x^k}{k!} f_N(x,t)dx, \qquad k = 0, 1,... \qquad (5.1.35)$$

Это выражение и определяет вероятность образования заданного количества событий для дважды стохастического пуассоновского процесса.

Распределение (35) впервые было введено Коксом (1955) в связи с задачами прикладной статистики; в физической литературе, имеющей дело с распределением отсчетов при регистрации излучения детекторами, оно именуется распределением Мандела (1958, 1959). Мы будем называть выражение (35) *распределением Кокса-Мандела.*

Для вычисления моментов величины $\mathfrak{N}(t)$ - числа событий в интервале времени длины t - целесообразно определить производящую функцию распределения Кокса-Мандела. Из (П.4.3) и (35) следует

$$G(s,t) = \int_0^\infty e^{-x(1-s)} f_N(x,t)\,dx, \qquad (5.1.36)$$

т.е. производящая функция представляет собой преобразование Лапласа плотности распределения величины $N(t)$. Из (П.4.4) и (36) находим:

$$M(t) \equiv \langle \mathfrak{N}(t) \rangle = \int_0^\infty x f_N(x,t)\,dx = \langle N(t) \rangle = \bar{n}t \qquad (5.1.37)$$

- среднее число событий линейно растет с увеличением длины интервала времени, в течение которого ведется наблюдение. Для дисперсии $D_{\mathfrak{N}}(t)$ величины $\mathfrak{N}(t)$ те же формулы дают

$$D_{\mathfrak{N}}(t) = \langle N^2(t) \rangle + \langle N(t) \rangle [1 - \langle N(t) \rangle] =$$

$$= \langle N(t) \rangle + D_N(t), \qquad (5.1.38)$$

где $D_N(t)$ - дисперсия $N(t)$. Принимая во внимание (37), запишем последний результат в виде

$$D_{\mathfrak{N}}(t) = M(t) + D_N(t). \qquad (5.1.39)$$

Здесь первое слагаемое в правой части представляет собой обычный *пуассоновский шум* - дисперсию количества импульсов при постоянной средней интенсивности. Стохастичность $n(t)$ вводит дополнительное слагаемое, так что итоговая дисперсия равна сумме пуассоновского шума и дисперсии непрерывной случайной функции $N(t)$.

С помощью (33) и (34) выразим полную дисперсию через ав-

токовариационную функцию $n(t)$:

$$D_{\mathfrak{N}}(t) = \bar{n}t + 2\int_0^t (t-\tau)\ \psi_n(\tau)d\tau. \qquad (5.1.40)$$

Индекс дисперсии дважды стохастического пуассоновского процесса равен

$$v_{\mathfrak{N}}(t) = 1 + \frac{2}{\bar{n}t}\int_0^t (t-\tau)\psi_n(\tau)d\tau. \qquad (5.1.41)$$

Согласно теореме Винера-Хинчина (1.5.13), автоковариационной функции колебаний средней интенсивности $\psi_n(\tau)$ отвечает спектральная плотность $g_n(\nu)$, содержащая информацию об амплитудах колебаний $n(t)$ всевозможных характерных периодов. Именно $g_n(\nu)$ и нужно обычно оценить на основании экспериментальных данных - совокупности моментов регистрации событий $\{t_k\}$. Функция $g_n(\nu)$ и индекс дисперсии процесса $\mathfrak{N}(t)$ связаны соотношением

$$v_{\mathfrak{N}}(t) = 1 + \frac{t}{\bar{n}}\int_{-\infty}^{\infty} \mathrm{sinc}^2(\nu t)g_n(\nu)d\nu, \qquad (5.1.42)$$

следующим из (41) и (1.5.13).

Пример 5.1.4. Предположим, что автоковариационная функция колебаний средней интенсивности имеет вид:

$$\psi_n(\tau) = \sigma_n^2\, e^{-|\tau|/\tau_n}, \quad -\infty < \tau < \infty\ . \qquad (5.1.43)$$

Для этого процесса индекс дисперсии

$$v_{\mathfrak{N}}(t) = 1 + \frac{2\tau_n\sigma_n^2}{\bar{n}} \cdot \frac{e^{-u}-1+u}{u}\ , \qquad (5.1.44)$$

где $u = t/\tau_n$. Функция $(e^{-u}-1+u)/u$ монотонно возрастает от 0 до 1 с ростом u, и при $t \gg \tau_n$ индекс дисперсии постоянен. Спектральная плотность находится путем подстановки (43) в (1.5.13):

$$g_n(\nu) = \frac{2\tau_n\sigma_n^2}{1+(2\pi\tau_n\nu)^2}\ . \qquad (5.1.45)$$

Задавая не только $\psi_n(\tau)$, но и функцию распределения средней интенсивности, можно углубить описание процесса.

Не составляет труда предсказать общий характер поведения

$D_{\text{Л}}(t)$ или $v_{\text{Л}}(t)$ при малых и больших значениях t. Предположим, что колебания $n(t)$ происходят с некоторым характерным временем τ_n, которое определим соотношениями

$$\tau_n = \frac{1}{B}\int_0^\infty \tau\psi_n(\tau)d\tau, \qquad B = \int_0^\infty \psi_n(\tau)d\tau. \tag{5.1.46}$$

Тогда при $t \ll \tau_n$ в (41) можно считать $\psi_n(\tau) \simeq \psi_n(0) = \sigma_n^2$. В противоположном случае, когда $t \gg \tau_n$, верхний предел интегрирования в (41) можно принять бесконечно большим, так как в области $\tau \gg \tau_n$ функция $\psi_n(\tau)$ очень мала. В итоге получаем

$$v_{\text{Л}}(t) \simeq \begin{cases} 1 + \dfrac{\sigma_n^2}{\bar{n}}\ t, & t \ll \tau_n, \\ 1 + \dfrac{2B}{\bar{n}}\left(1 - \dfrac{\tau_n}{t}\right), & t \gg \tau_n \end{cases} \tag{5.1.47}$$

(ср. с результатами примера 5.1.4). Эти равенства лишь наличием постоянного слагаемого, отвечающего пуассоновскому шуму, отличаются от соответствующих равенств теории броуновского движения (см., например, Ахманов и др., 1981, гл. 2, § 6), где постоянная B называется *коэффициентом диффузии*. Как отмечалось в § 3.6, пройденный броуновской частицей путь есть интеграл от хаотического процесса, близкого к белому шуму. Согласно (32), $N(t)$ также является интегралом от случайной интенсивности процесса $n(t)$. Дисперсия интегральных величин пропорциональна t^2, пока сказывается инерционность между отдельными случайными «толчками» (т.е. при $t \ll \tau_n$), а затем возрастает по диффузионному закону $D(t) \propto t$, отражающему результат многих таких «толчков» (при $t \gg \tau_n$).

Мы остановились на свойствах дисперсии дважды стохастического процесса Пуассона по той причине, что именно ее поведение в области $t \sim \tau_n$ позволяет, в принципе, найти из наблюдений характерное время переменности τ_n, а затем и другие параметры флуктуаций средней интенсивности (§§ 5.3, 5.4).

§ 5.2. Спектральный анализ импульсных последовательностей

Исследователь, анализирующий реальную совокупность то-

чечных событий, в лучшем случае может иметь лишь косвенные соображения относительно внутренней структуры этой последовательности. Так, мы увидим, в § 5.5, что существуют теоретические основания ожидать близости последовательности импульсов на выходе детектора оптического излучения к однородному пуассоновскому процессу, если время экспозиции существенно превосходит время когерентности излучения. На практике, однако, фотоумножитель и тракт регистрации вносят искажения, которые можно выявить только путем изучения достаточно длинных временных рядов (§ 5.6). В других случаях можно подозревать наличие тренда или периодичности в совокупности случайных отсчетов. Чаще всего ожидаемые отличия от пуассоновского процесса невелики, и потому необходимо сначала 1) проверить соответствие экспериментальных данных «совершенно случайному» процессу - однородному процессу Пуассона; 2) при наличии значимых отклонений сопоставить данным наблюдений неоднородный или дважды стохастический пуассоновские процессы, выявить характерные амплитуду и временной масштаб колебаний средней интенсивности $n(t)$.

Принимая во внимание результаты, изложенные в предшествующих главах, можно сказать, что нас интересует оценка спектральной плотности величины $n(t)$, рассматриваемой либо как детерминированная, либо как случайная функция.

Вместе с тем следует подчеркнуть, что однородным пуассоновским процессом и его непосредственными обобщениями далеко не исчерпываются все виды случайных последовательностей, представляющие практический интерес. Например, часто хорошее описание экспериментальных данных дают так называемые *процессы восстановления*, для которых интервалы между событиями независимы, одинаково распределены, однако не подчиняются экпоненциальному распределению (5.1.14), как для однородного процесса Пуассона. В более общем случае последовательные интервалы могут быть зависимы.

Итак, представляется необходимым сначала дать несколько более общее описание точечных случайных процессов с тем, чтобы далее на основании наблюдаемой совокупности $\{t_k\}$ моментов наступления событий получить надежные оценки теоретических характеристик. Этот подход и рассматривается в

данном параграфе. Более подробное изложение и строгие доказательства основных результатов имеются в книге Кокса и Льюиса (1966). Пример спектрального анализа импульсной случайной последовательности приведен в § 5.4.

Стационарные точечные процессы можно рассматривать как обобщение пуассоновского процесса, при котором сохраняется лишь свойство ординарности событий. Определим процесс, задавая последовательность, вообще говоря, зависимых интервалов между событиями $\{X_k = t_{k+1} - t_k\}$, $k = 1, 2, \ldots$, имеющих одну и ту же плотность распределения $f_X(\tau)$, $0 \leqslant \tau < \infty$. Следует заметить, что статистические свойства последовательности $\{X_k\}$ отличаются от свойств последовательности интервалов $\{Y_k\}$, образованной при случайном выборе начального момента наблюдения t_0. Дело в том, что произвольный выбор t_0 помещает эту точку на совсем не типичный интервал между событиями, и память об этом сохраняется достаточно долго. Отсылая за детальным обсуждением этого вопроса к книгам Феллера (1966) и Кокса и Льюиса (1966), ограничимся рассмотрением свойств совокупности $\{X_k\}$.

Для того чтобы стала яснее связь излагаемого ниже аппарата с материалом предшествующих глав, представим себе, что последовательные реализации свободных интервалов между событиями X_1, $X_2, \ldots$ расположены в направлении оси ординат, а по оси абсцисс откладываются их номера k. Мы имеем тогда обыкновенный *равномерный ряд*, к которому можно применить стандартные методы анализа, развитые выше. В частности, введем спектральную плотность длин интервалов $g_X(z)$, и если, скажем, эта функция характеризуется выделяющимся максимумом на некоторой (безразмерной) частоте $z_0 \in [-1/2, 1/2]$, то не исключена периодичность процесса с характерным количеством интервалов на периоде $|z_0|^{-1}$ или периодом $P_0 \simeq \simeq \langle X \rangle / |z_0|$, где $\langle X \rangle$ - средняя длина свободного интервала времени.

Определим прежде всего по аналогии с (2.3.1) автоковариационную функцию последовательности свободных интервалов соотношением

$$\psi_k = \langle X_1 X_{k+1} \rangle - \langle X \rangle^2, \quad \psi_{-k} = \psi_k, \quad k = 0, \pm 1, \ldots, \qquad (5.2.1)$$

так что $\psi_0 = D_X$ - дисперсия X. Далее обычным образом (ср. (1.5.13) и (2.3.2)) сопоставим автоковариационной функции спектральную плотность

$$g_X(z) = \sum_{k=-\infty}^{\infty} e^{-i\cdot 2\pi zk}\,\psi_k =$$

$$= D_X\left[1 + 2\sum_{k=1}^{\infty}\cos(2\pi zk)\rho_k\right], \quad -\frac{1}{2} \leqslant z \leqslant \frac{1}{2}, \qquad (5.2.2)$$

где $\rho_k = \psi_k/D_X$ коэффициенты корреляции $\{X_k\}$. Обращение (2) имеет вид

$$\psi_k = \int_{-1/2}^{1/2} e^{i\cdot 2\pi zk}\, g_X(z)\,dz, \quad k = 0, \pm 1, \ldots \qquad (5.2.3)$$

Последняя формула включает в себя и аналог теоремы Рэлея:

$$\psi_0 = D_X = \int_{-1/2}^{1/2} g_X(z)\,dz\,. \qquad (5.2.4)$$

Приведенные соотношения дают описание точечного процесса посредством задания свойств совокупности $\{X_k\}$ - свободных промежутков времени между событиями. Другой подход связан с изучением вероятностного распределения $\{p_k(t)\}$ величины $\mathfrak{N}(t)$ - количества событий, происшедших за время t, начиная с произвольно выбранного начала отсчета времени (но не с произвольно выбранного события!). Теоретически оба подхода эквивалентны, однако на практике изучаются лишь первые моменты характеристик процесса, и тогда указанные его описания дают взаимно дополнительную информацию о механизме генерации случайных импульсов.

Среднее число событий в промежутке $(0,t]$ и средняя длина свободного интервала связаны соотношениями

$$\langle\mathfrak{N}(t)\rangle \equiv M(t) = \overline{n}t, \qquad (5.2.5)$$

$$\overline{n} = \langle X\rangle^{-1}, \qquad (5.2.6)$$

где, очевидно, $\overline{n}$ - средняя интенсивность процесса. Далее, автоковариационная функция интервалов и распределение числа событий удовлетворяют следующим равенствам:

$$\psi_0 = D_X = \langle X\rangle\left[2\int_0^\infty p_0(t)\,dt - \langle X\rangle\right],$$
$$\psi_k = \langle X\rangle\left[\int_0^\infty p_k(t)\,dt - \langle X\rangle\right], \quad k = 1, 2, \ldots \tag{5.2.7}$$

функция распределения длины свободного интервала $F_X(\tau)$ и вероятность отсутствия событий в $(0,t]$ связаны уравнением

$$p_0(t) = 1 - \frac{1}{\langle X\rangle}\int_0^t\left[1 - F_X(\tau)\right]d\tau, \tag{5.2.8}$$

решение которого относительно плотности $f_X(\tau)$ дает

$$f_X(\tau) = F_X'(\tau) = \langle X\rangle\,\frac{d^2 p_0(t)}{dt^2}. \tag{5.2.9}$$

Таким образом, знание $\{p_k(t)\}$ позволяет без труда найти по формулам (6) - (9) основные характеристики распределения интервалов между событиями.

Поскольку точечный процесс характеризуется как совокупностью $\{X_k\}$, так и накопленным числом событий $\mathfrak{N}(t)$, желательно ввести наряду со спектральной плотностью $g_X(z)$ и спектральную плотность целочисленного процесса $\mathfrak{N}(t)$. Нужно, однако, учесть, что $\mathfrak{N}(t)$ не является стационарным процессом, как мы его определили в § 1.5, - среднее значение (5) линейно растет со временем. Стационарный случайный процесс образует лишь приращение $\Delta\mathfrak{N}(t)$ за фиксированный промежуток времени. Строгий подход к определению автоковариационной функции приращений $\psi_{\mathfrak{N}}(\tau)$ изложен Коксом и Льюисом (1966); ограничиваясь соображениями на интуитивном уровне, положим

$$\psi_{\mathfrak{N}}(\tau) = \left\langle \frac{d\mathfrak{N}(t)}{dt}\cdot\frac{d\mathfrak{N}(t+\tau)}{dt}\right\rangle - \left\langle\frac{d\mathfrak{N}(t)}{dt}\right\rangle\left\langle\frac{d\mathfrak{N}(t+\tau)}{dt}\right\rangle \tag{5.2.10}$$

и, соответственно теореме Винера-Хинчина,

$$g_{\mathfrak{N}}(\nu) = \int_{-\infty}^{\infty} e^{-i\cdot 2\pi\nu\tau}\psi_{\mathfrak{N}}(\tau)\,d\tau. \tag{5.2.11}$$

Функцию $g_{\mathfrak{N}}(\nu)$ называют спектральной плотностью целочисленного процесса приращений $d\mathfrak{N}(t)$.

Рассмотрим $g_{\mathfrak{N}}(\nu)$ для дважды стохастического процесса Пуассона. Проведем в (10) усреднение в два этапа: сначала по числу событий при фиксированной реализации средней интенсивности $n(t)$, а затем по ансамблю случайных функций

$\{n(t)\}$. Ввиду (5.1.18) три последовательных средних в правой части (10) на первом этапе оказываются равными $n(t)n(t+\tau)$, $n(t)$ и $n(t+\tau)$ соответственно. Проводя второе усреднение и сравнивая результат с (5.1.31), получаем:

$$\phi_{\mathfrak{N}}(\tau) = \langle n(t)n(t+\tau)\rangle - (\bar{n})^2 = \phi_n(\tau) \qquad (5.2.12)$$

- автоковариационная функция процесса $\mathfrak{N}(t)$ и аналогичная функция средней интенсивности $n(t)$ для дважды стохастического пуассоновского процесса равны. Отсюда следует, что совпадают и спектральные плотности:

$$g_{\mathfrak{N}}(\nu) = g_n(\nu). \qquad (5.2.13)$$

Как мы увидим ниже, этот важный результат позволяет получить простую оценку $g_{\mathfrak{N}}(\nu)$ для дважды стохастического процесса Пуассона.

Оценки параметров и спектральных плотностей. Рассмотрим наиболее важный для практики этап - сопоставление теоретическим характеристикам импульсного процесса некоторых оценок, опирающихся только на зарегистрированную последовательность моментов $\{t_k\}$ или связанную с ней последовательность свободных интервалов $\{X_k = t_{k+1} - t_k\}$, $k = 1, 2, \ldots$

Если на опыте зарегистрировано N событий (рис. 5.1.1) и соответственно $N - 1$ интервалов $(X_1, X_2, \ldots, X_{N-1})$, то оценкой средней длины $\langle X\rangle$ может служить величина

$$\overline{X} = \frac{1}{N-1}\sum_{k=1}^{N-1} X_k. \qquad (5.2.14)$$

Она является *несмещенной* (т.е. $\langle\overline{X}\rangle = \langle X\rangle$) и *состоятельной* (т.е. $D_{\overline{X}} \xrightarrow[N\to\infty]{} 0$) оценкой истинного среднего. Аналогично для целочисленного процесса $\mathfrak{N}(t)$ несмещенной и состоятельной оценкой средней интенсивности событий является

$$\overline{n} = \frac{N}{T} \simeq \frac{1}{\overline{X}}, \qquad (5.2.15)$$

где $T = t_N - t_0$ - продолжительность наблюдений.

Нетрудно получить и оценку спектральной плотности интервалов между событиями. Действительно, последовательность

$\{X_k\}$ представляет собой обычный равномерный временной ряд, так что естественно образовать периодограмму типа (2.1.22):

$$P_X(z) = \left| \frac{1}{N-1} \sum_{k=1}^{N-1} e^{-i \cdot 2\pi z k} (X_k - \bar{X}) \right|^2, -\frac{1}{2} \leqslant z \leqslant \frac{1}{2}, \qquad (5.2.16)$$

и по аналогии с (2.3.7) рассматривать величину

$$\hat{g}_X(z) \equiv NP_X(z), \quad z \neq 0; \pm 0{,}5 \qquad (5.2.17)$$

как оценку истинной спектральной плотности (2).

Для выяснения значимости наблюдаемых максимумов периодограммы в § 2.3 предполагалось, что референтный процесс - нормальный с дисперсией σ^2; при этом было показано, что $P_N(\nu)$ экспоненциально распределена с одинаковыми средним значением и стандартным отклонением, равными σ^2/N. В случае импульсного процесса референтным является пуассоновский процесс, для которого $\{X_k\}$ распределены *экспоненциально* с равными средним значением и стандартным отклонением: $\langle X \rangle = D_X^{1/2} = 1/n$ (см. формулы (5.1.16) и (5.1.17)). Тем не менее вследствие центральной предельной теоремы при больших значениях N периодограмма $P_X(z)$ по-прежнему подчиняется экспоненциальному распределению, причем

$$\langle P_X(z) \rangle = D^{1/2}\left[P_X(z)\right] \simeq \frac{D_X}{N} = \frac{1}{Nn^2}. \qquad (5.2.18)$$

Заменяя здесь n экспериментальной оценкой (15), получаем, что величина

$$\eta \equiv \frac{NP_X(z)}{(\bar{X})^2} = \frac{\hat{g}_X(z)}{(\bar{X})^2} \qquad (5.2.19)$$

подчиняется экспоненциальному распределению с плотностью e^{-x}, $0 \leqslant x < \infty$. Это обстоятельство позволяет перенести все результаты § 2.3 и на случай оценивания значимости периодограммы ряда свободных интервалов $\{X_k\}$. В частности, максимальное значение η на интервале $0 \leqslant z \leqslant 1/2$ подчиняется распределению Уокера (2.3.31) и все последующие расчеты уровня значимости сохраняют справедливость.

Заметим наконец, что $P_X(z)$, как и всякая периодограмма, характеризуется значительными флуктуациями, не уменьшающи-

мися при $N \to \infty$, т.е. (17) является *несостоятельной* оценкой истинной спектральной плотности. Сглаживание $P_X(z)$ и получение состоятельной оценки может быть проведено при помощи методов, изложенных в § 2.4. Возможно и непосредственное применение метода максимума энтропии в том виде, как он представлен в § 2.5.

Обратимся к нахождению оценки спектральной плотности целочисленного процесса $g_{\mathfrak{N}}(\nu)$, играющей основную роль в приложениях.

Рассмотрим сначала дважды стохастический пуассоновский процесс, для которого выполняется соотношение (13), так что оцениванию подлежит функция $g_n(\nu)$. Предположим, что процесс наблюдался в течение интервала времени $(0,T]$ и $\tilde{n}(t)$ - реализация средней интенсивности. Согласно определению спектральной плотности случайного процесса, данному формулами (1.5.7), (1.5.11) и (1.5.12), мы имеем:

$$g_n(\nu) = \lim_{T\to\infty} g_n^{(T)}(\nu), \tag{5.2.20}$$

где

$$g_n^{(T)}(\nu) = \frac{1}{T}\left|\int_0^T e^{-i\cdot 2\pi\nu t}\,\tilde{n}(t)\,dt\right|^2. \tag{5.2.21}$$

Представим далее $\tilde{n}(t)$ в виде суммы δ-функций Дирака:

$$\tilde{n}(t) = \sum_{k=1}^{N} \delta(t-t_k), \tag{5.2.22}$$

где $\{t_k\}$ - последовательность моментов регистрации событий. Основанием для представления (22) служит тот факт, что процесс приращений испытывает скачки при $t = t_k$; с другой стороны, интегрирование (22) по некоторому промежутку $[t',t'']$ дает ввиду (П.2.1) действительное количество событий, случившихся в этом промежутке. Подстановка (22) в (21) приводит к выражению

$$g_n^{(T)}(\nu) = \frac{1}{T}\left|\sum_{k=1}^{N} e^{-i\cdot 2\pi\nu t_k}\right|^2 \equiv \frac{N}{T}\, I_N(\nu), \tag{5.2.23}$$

где безразмерная функция

$$I_N(\nu) = \frac{1}{N}\left|\sum_{k=1}^{N} e^{-i\cdot 2\pi\nu t_k}\right|^2, \quad 0 \leqslant \nu < \infty. \tag{5.2.24}$$

Ввиду (13) мы можем представить оценку спектральной плотности целочисленного дважды стохастического процесса в следующем виде:

$$\hat{g}_{\mathfrak{N}}(\nu) = \frac{N}{T} I_N(\nu) \simeq I_N(\nu)/\overline{X}. \qquad (5.2.25)$$

Соотношения (24) и (25) являются искомыми. Хотя они были выведены для дважды стохастического процесса Пуассона, можно показать, что (25) сохраняет справедливость при оценивании спектральной плотности *произвольного* стационарного точечного процесса (Кокс и Льюис, 1966, § 5.5). Величина $2I_N$ совпадает со *статистикой Рэлея* (1880), свойства которой обсуждаются Мардиа (1972), Чандрасекаром (1943) и - в контексте счета фотонов - Гибсоном и др. (1982), Лихи и др. (1983 b). Следуя Фраске (1985), мы будем называть функцию $I_N(\nu)$ *импульсным спектром* процесса. С точностью до множителя N импульсный спектр получается при замене всех отсчетов $\{y_k\}$ в обычном определении периодограммы (1.6.10), (1.6.11) единичными отсчетами $y_k \equiv 1$ и пропорционален, таким образом, плотности спектрального окна данных (1.6.19):

$$I_N(\nu) = N\,|W_N(\nu)|^2. \qquad (5.2.26)$$

Поскольку $I_N(0) = N$, точка $\nu = 0$ является выделенной в том смысле, что в ней импульсный спектр неограниченно растет с увеличением количества отсчетов. Полагая $N \gg 1$ и отсчитывая время от середины интервала наблюдения, мы имеем для пуассоновского набора $\{t_k\}$

$$\sum_{k=1}^{N} e^{-i\cdot 2\pi\nu t_k} \simeq N\langle e^{-i\cdot 2\pi\nu t_k}\rangle \simeq \frac{N}{T}\int_{-T/2}^{T/2} e^{-i\cdot 2\pi\nu t}\,dt =$$

$$= N\frac{\sin(\pi\nu T)}{\pi\nu T} = N\,\mathrm{sinc}(\nu T), \qquad (5.2.27)$$

так что определение (24) импульсного спектра можно несколько изменить следующим образом:

$$\tilde{I}_N(\nu) = \frac{1}{N}\left|\sum_{k=1}^{N} e^{-i\cdot 2\pi\nu t_k} - N\,\mathrm{sinc}(\nu T)\right|^2, \quad 0 \leqslant \nu < \infty. \qquad (5.2.28)$$

При этом $\tilde{I}_N(0) = 0$, а в точках, где νT равно целому числу, обычный и модифицированный импульсные спектры совпадают. В

случае $N \gg 1$ различие спектров фактически проявляется лишь для начала координат; численные расчеты можно начинать со значения частоты $\nu_{min} = 1/T$.

Учитывая связь импульсного спектра с пеpодиограммой конкретного ряда $\{y_k \equiv 1\}$, нетрудно предсказать статистические свойства $I_N(\nu)$. Именно, для случайных равномерно распределенных на промежутке наблюдения моментов $\{t_k\}$ величина $I_N(\nu)$ должна быть экспоненциально распределена (см. § 2.3) с равными между собой средним значением и стандартным отклонением. Как и в § 2.3, рассмотрим дискретный набор частот $\nu_j = j/T$, где $j \neq 0$ - целые числа. Тогда из (3.2.5) и (3.2.18) следует

$$\langle I_N(\nu_j) \rangle = D^{1/2}[I_N(\nu_j)] = NR_N(\nu_j) = 1 \qquad (5.2.29)$$

- величина $\eta_j \equiv I_N(\nu_j) = \tilde{I}_N(\nu_j)$ подчиняется экспоненциальному распределению с плотностью $p_\eta(x) = e^{-x}$, $0 \leqslant x < \infty$. Дальнейшая методика оценивания значимости максимумов импульсного спектра изложена в § 2.3. Следует однако подчеркнуть, что ввиду неравномерности ряда $\{t_k\}$ импульсный спектр не является периодической функцией частоты, а потому нужно разумным образом ограничить интервал частот, в пределах которого оценивается спектральная плотность.

Ввиду (25) и (29) первые моменты

$$\langle \hat{g}_\eta(\nu_j) \rangle = D^{1/2}\left[\hat{g}_\eta(\nu_j)\right] = \frac{N}{T} \simeq 1/\overline{X}, \qquad (5.2.30)$$

и мы могли бы рассматривать в качестве η_j соответствующее отношение спектральных плотностей. Причины «изрезанности» импульсного спектра и методы сглаживания оценок неоднократно обсуждались выше.

Импульсные последовательности обычно реже страдают от пропусков наблюдательного времени, чем ряды наблюдений нестационарных звезд. Если, тем не менее, пропуски имели место, их можно характеризовать функцией

$$\lambda(t) = \begin{cases} 1 \text{ на интервалах, где наблюдения велись,} \\ 0 \text{ в остальных случаях,} \end{cases} \qquad (5.2.31)$$

а затем следующим образом корректировать импульсный спектр

(Фраска, 1985):

$$I_N(\nu) = \frac{1}{N}\left| \sum_{k=1}^{N} e^{-i\cdot 2\pi\nu t_k} - N\,\frac{\Lambda(\nu)}{\Lambda(0)}\right|^2, \qquad (5.2.32)$$

где

$$\Lambda(\nu) = \int_{-T/2}^{T/2} e^{-i\cdot 2\pi\nu t}\lambda(t)\,dt. \qquad (5.2.33)$$

Для случая когда $\lambda(t) \equiv 1$ в пределах интервала наблюдения, мы возвращаемся к определению (28).

Итак, статистические свойства последовательности N импульсов, зарегистрированных в моменты $\{t_k\}$, можно охарактеризовать либо оценкой $\hat{g}_X(z)$ спектральной плотности интервалов между событиями (17), либо импульсным спектром (24). При этом для «совершенно случайного» - пуассоновского процесса с постоянной интенсивностью средние значения $\langle\hat{g}_X(z)\rangle = (\overline{X})^2$, $\langle I_N(\nu)\rangle = 1$, а отношения частных значений к средним распределены экспоненциально с плотностью e^{-x}, $0 \leqslant x < \infty$. Значимые отклонения от средних указывают на возможные колебания интенсивности событий с выделенными периодами.

Что касается сравнительной эффективности двух обсуждаемых методов при обнаружении периодических колебаний, то предпочтение, как нам кажется, следует отдать импульсному спектру. Дело в том, что порядковый номер интервала в последовательности $\{X_k\}$ связан со временем не прямым, а сложным стохастическим образом, и эта двойная стохастичность - самих длин интервалов и их номеров - сильно снижает чувствительность метода, основанного на анализе $g_X(z)$.

Быстрая переменность. Изучение переменности с характерным временем порядка среднего интервала между событиями $\overline{X}$ имеет свои особенности. Здесь экономнее регистрировать интервалы между событиями $\{X_k\}$, а не количество событий, происшедших в пределах малых интервалов длительности Δt: в случае $\Delta t < \overline{X}$ последовательность $\mathfrak{N}(\Delta t)$ состоит в основном из нулей. С теоретической точки зрения речь идет о поведении автоковариационной функции свободных интервалов (1) при малых k или о поведении спектральной плотности $g_{\mathfrak{N}}(\nu)$ в области частот $\nu \geqslant (\overline{X})^{-1} = \overline{n}$.

Если вся последовательность содержит N импульсов, то минимальный промежуток времени между соседними событиями порядка $\overline{X}/N$, и частоты выше $\sim N\overline{n}$ принципиально недоступны. С другой стороны, согласно теореме отсчетов (§ 2.2) для изучения переменности с характерным временным масштабом τ этот интервал должен содержать не менее двух отсчетов. Принимая в качестве критической длительности интервала $\tau_* = 2\overline{X}$, находим, что аналогом критической частоты Найквиста для импульсных последовательностей является $\nu_* = (2\overline{X})^{-1}$. Таким образом, проблема заключается в восстановлении свойств импульсного процесса в диапазоне частот с приблизительными граничными значениями

$$\nu_* = \overline{n}/2, \qquad \nu_{max} = N\overline{n} . \tag{5.2.34}$$

Ввиду стохастического характера последовательности импульсов ν_* обладает не всеми свойствами найквистовского предела - этот предел в данном случае «размыт» вблизи ν_*.

Эквивалентный способ задания условий быстрой переменности связан с использованием введенного в § 5.1 параметра $\gamma(\nu) \equiv \overline{n}P = \overline{n}/\nu$ - среднего количества событий за исследуемый (пробный) период. Ввиду (34) область быстрой переменности определяется неравенством $\gamma \leqslant 2$.

Если речь идет о выделении периодических компонентов колебаний средней интенсивности $n(t)$, то нет оснований ожидать существенной потери эффективности импульсного спектра в области $\gamma < 1$. Результаты численных расчетов, описанные в § 5.4, подтверждают этот вывод.

Считая процесс регистрации фотонов пуассоновским с переменной средней интенсивностью $n(t)$, Шварцман (1977) предложил с целью изучения быстрой переменности использовать выражение (5.1.23) для плотности распределения $\bar{f}(\tau)$ длин свободных интервалов между событиями. Величина $\bar{f}(\tau)/(\overline{n}e^{-\overline{n}\tau}) - 1$, т.е. с точностью до постоянного слагаемого плотность, нормированная на ожидаемую при постоянной интенсивности экспоненциальную плотность распределения (5.1.14), была названа им y_2-функцией. Поскольку при этом производится усреднение по большому интервалу наблюдения, фазовая упорядоченность колебаний $n(t)$ нивелируется, и указанное

отношение нечувствительно к *периодическим* изменениям интенсивности (см. также численный пример в § 5.4). Менее очевидным представляется вопрос об эффективности использования усредненного распределения $\{X_k\}$ для обнаружения *непериодических* вариаций $n(t)$. Вопрос этот требует специального изучения. Из общих соображений можно ожидать, что интегральные характеристики, а $\bar{f}(\tau)$ является именно таковой, не относятся к числу наиболее эффективных. Наконец, при *стохастической* переменности $n(t)$ уже нельзя опираться на выражение (5.1.23), а нужно обратиться к формуле (9), конкретизированной в соответствии с предположениями о характере рассматриваемого процесса.

Оценка спектральной плотности методом максимального правдоподобия. Высокая эффективность информационных критериев восстановления изображения и спектра (см. §§ 2.5, 3.8) позволяет предположить, что применение метода максимума правдоподобия и в случае импульсных процессов приведет к оптимальной оценке спектральной плотности. Ограничимся здесь краткими замечаниями.

Пусть на интервале времени $[0,T]$ зарегистрированы N импульсов в моменты $(t_1, t_2, \ldots, t_N)$. Предположим, что эта совокупность является реализацией неоднородного пуассоновского процесса, средняя интенсивность которого $n(t)$ представлена в виде интеграла Фурье:

$$n(t) = \int_{-\infty}^{\infty} e^{i\cdot 2\pi\nu t} N(\nu)\,d\nu. \tag{5.2.35}$$

Вероятность получения наблюдаемой реализации пропорциональна величине

$$P = \exp\left[-\int_0^T n(t)\,dt\right] \prod_{k=1}^{N} n(t_k). \tag{5.2.36}$$

Необходимым условием экстремума $\ln P$ относительно вариаций $N(\nu)$ является выполнение уравнений

$$\sum_{k=1}^{N} \frac{e^{i\cdot 2\pi\nu t_k}}{n(t_k)} = \frac{e^{i\cdot 2\pi\nu T}-1}{i\cdot 2\pi\nu},$$

$$n(t_k) = \int_{-\infty}^{\infty} e^{i\cdot 2\pi\nu t_k} N(\nu)\,d\nu, \tag{5.2.37}$$

решение которых дает оценку максимального правдоподобия $\hat{N}(\nu)$, а затем и аналогичную оценку спектральной плотности $\hat{G}(\nu) = |\hat{N}(\nu)|^2$.

§ 5.3. Методы, опирающиеся на группировку событий

При анализе обычных временных рядов в предыдущих главах использовались два основных подхода: дискретное преобразование Фурье и непараметрические методы, связанные с группированием данных по фазовым интервалам. Аналогичные подходы применимы и в анализе импульсных последовательностей. Рассмотренный в § 5.2 импульсный спектр $I_N(\nu)$ представляет собой оценку спектральной плотности Фурье точечного процесса. В настоящем параграфе мы обратимся к группе методов, использующих разбиение промежутка наблюдений длительности T на пробные системы интервалов. Именно эти методы наиболее употребительны в рентгеновской астрономии, тогда как при интерпретации лабораторных экспериментов нередко применяется импульсный спектр (Фраска, 1985).

Легко видеть, что характерные особенности двух указанных подходов сохраняются и при анализе импульсных случайных процессов. Так, построение спектра Фурье обеспечивает временное разрешение, присущее исходным данным, и не вносит элементов субъективности, обусловленных выбором биновой структуры. С другой стороны, при значительном количестве измерений, а именно этот случай реализуется чаще всего при счете фотонов, дисперсионные методы позволяют существенно сократить время обработки. Другим преимуществом непараметрических методов, которое неоднократно обсуждалось нами ранее, является их высокая чувствительность при несинусоидальной форме колебаний средней интенсивности. Описанные подходы связаны между собой, однако во многих отношениях они дают дополнительную информацию об изучаемом процессе.

Критерий χ^2 в различных модификациях применяется для решения многих задач статистики; в данном случае нас будет интересовать оценка значимости расхождения между некоторым наблюдаемым вероятностным распределением и наперед заданным (теоретическим) распределением.

Предположим, что весь промежуток изменения случайной величины разбит на r интервалов (бинов), содержащих соответственно $N_1, N_2, \ldots, N_r$ наблюденных значений, $\sum_1^r N_i = N$. Отношения $\{N_i/N\}$ представляют собой *наблюдаемые* частоты, отвечающие принятой биновой структуре. Интегрируя заданную плотность распределения $p(x)$ в пределах каждого бина, можно найти *теоретические* вероятности $\{p_i\}$ попадания в бины (как всегда, $\sum_1^r p_i = 1$). Вопрос заключается в следующем: можно ли считать, что отличие экспериментальных и теоретических частот обусловлено лишь неизбежными статистическими флуктуациями при образовании выборки конечного объема N из совокупности значений случайной величины с заданным теоретическим распределением? Для решения этого вопроса следует сопоставить наблюдаемое на опыте расхождение между $\{N_i/N\}$ и $\{p_i\}$ с типичными значениями такого расхождения, ожидаемыми при образовании выборок в N значений случайной величины с плотностью $p(x)$, т.е., как говорят, при справедливости *нулевой гипотезы* H_0.

Итак, прежде всего нужно ввести меру расхождения двух дискретных распределений вероятности. В качестве таковой естественно взять сумму $\sum_1^r c_i(N_i/N - p_i)^2$, где $\{c_i\}$ - весовые коэффициенты, подобранные надлежащим образом. Если, следуя Пирсону (1900), принять $c_i = N/p_i$, то получится мера расхождения

$$\chi^2 = \sum_{i=1}^{r} \frac{(N_i - Np_i)^2}{Np_i}, \qquad (5.3.1)$$

вероятностные свойства которой при справедливости H_0 и $N \gg 1$ не зависят от исходного распределения $p(x)$. Можно показать (см., например, Крамер, 1946, п.30.1), что это предельное распределение случайной величины (1) является χ^2-распределением с $r-1$ степенями свободы. Основные свойства χ^2-распределения перечислены в Приложении 4. В частности, среднее значение $\langle\chi^2\rangle = r-1$, дисперсия $D(\chi^2) \simeq 2(r-1)$, так что величина

$$\Omega = \frac{1}{\sqrt{2(r-1)}} \left[\sum_{1}^{r} \frac{(N_i - Np_i)^2}{Np_i} - (r-1) \right] \qquad (5.3.2)$$

имеет нулевое среднее и единичную дисперсию. При $r \gg 1$ можно считать, что Ω подчиняется гауссовскому распределению.

Как обычно при решении проблемы значимости, если наблюдаемое расхождение χ^2_{obs} существенно превосходит среднее значение $r-1$, то есть основания подозревать реальность отличия экспериментального $\{N_i/N\}$ и теоретического $\{p_i\}$ распределений. Степень расхождения характеризуется площадью «хвоста» вероятностного распределения $Pr(\chi^2 \geqslant \chi^2_{obs}) = q$, т.е. вероятностью q того, что чисто случайные факторы приведут к расхождению, не меньшему наблюдаемого в действительности. Если x_q - значение χ^2, отвечающее уровню значимости q, то этому же уровню отвечает значение Ω, равное $[x_q-(r-1)]/\sqrt{2(r-1)}$.

Выделение периодического компонента в импульсной последовательности методом сворачивания эпох проводится по сути дела так же, как описано в § 1.7 и гл. 4. Пусть $\{t_k\}$, $k = 1,\ldots, N$ - моменты регистрации событий, T - полное время наблюдений. Выберем пробный период $P \equiv \nu^{-1}$ и свернем последовательность с этим периодом. Согласно теореме Пуанкаре (1912; см. также Феллер, 1966, Гибсон и др., 1982), случайный временной ряд достаточной длины, свернутый с произвольным периодом, образует равномерное фазовое распределение. Если же значение P лежит вблизи реального периода P_0 колебания средней интенсивности поступления событий, то распределение отсчетов на фазовой диаграмме обнаружит систематическое изменение. Другими словами, выбранная подходящим образом мера упорядоченности Γ_P изменения фазовой плотности числа отсчетов будет иметь экстремум вблизи P_0.

Согласно изложенному в §§ 1.6 и 2.3, пробные периоды должны подбираться так, чтобы шаг по частоте $\delta\nu$ был равномерным; разумно принять $\delta\nu \simeq (sT)^{-1} - (10T)^{-1}$. В этой связи удобнее рассматривать зависимость Γ не от пробного периода P, а от частоты ν. Как и ранее, функцию $\Gamma(\nu)$ будем называть

периодограммой.

Итак, пусть t_0 - произвольное начало отсчета времени, $FR(z)$ означает дробную часть z и

$$x_k(\nu) = FR[\nu(t_k - t_0)], \quad k = 1, 2, \ldots, N, \qquad (5.3.3)$$

- значения фазы, соответствующие моментам $\{t_k\}$ для частоты ν. Наша задача заключается в том, чтобы задать по возможности более эффективную меру упорядоченности $\{x_k\}$. Один из наиболее простых способов заключается в том, чтобы проверить равномерность распределения $\{x_k\}$. Стандартной мерой отклонения от заданного распределения является величина χ^2, определенная формулой (1), где под $\{p_i\}$ теперь нужно понимать значения $\{\Delta x_i\}$ ширин бинов фазового промежутка $[0,1)$:

$$\chi^2(\nu) = \sum_{i=1}^{r} \frac{(N_i - N\Delta x_i)^2}{N\Delta x_i} . \qquad (5.3.4)$$

Количество бинов r выбирается так, чтобы содержимое каждого из них составляло не менее 10. Если ΔT_i - время наблюдения, соответствующее i-му бину, то $\Delta x_i = \Delta T_i/T$ и

$$\chi^2(\nu) = \sum_{i=1}^{r} \frac{(N_i - \bar{n}\Delta T_i)^2}{\bar{n}\Delta T_i} , \qquad (5.3.5)$$

где $\bar{n} = N/T$ - оценка средней интенсивности событий. В форме (5) критерий χ^2 применим и при наличии пропусков наблюдательного времени.

Как указывалось в предыдущем пункте, для далеких от ν_0 значений ν величина (5) подчиняется χ^2-распределению с $r-1$ степенями свободы; ее среднее значение $\langle\chi^2\rangle = r-1$. Отклик $\chi^2(\nu_0)$ вблизи частоты истинной периодичности зависит от формы и амплитуды периодического компонента. Так, мы покажем в § 5.4, что для гармонического изменения средней интенсивности с относительной амплитудой $\delta n/\bar{n} = \mu$ максимальный отклик $\chi^2(\nu_0) \simeq N\mu^2/2$. Оценка значимости наблюдаемых выбросов периодограммы $\chi^2(\nu)$ должна опираться на общепринятую методику (см., например, Лихи и др., 1983 *а*); при этом следует учитывать, что значения $\chi^2(\nu)$, разделенные интерва-

лом частот $\sim T^{-1}$, независимы друг от друга (§ 2.3).

Некоторой субъективности, обусловленной выбором биновой структуры, можно избежать, усредняя $\chi^2(\nu)$ по различным системам бинов. Другой путь заключается в выборе *поточечной* характеристики меры упорядоченности $\{x_k\}$. В качестве таковой можно, например, принять

$$E(\nu) = 1 + \frac{1}{\log N} \sum_{k=1}^{N} \Delta x_k \log \Delta x_k, \qquad (5.3.6)$$

где интервалы между фазовыми отсчетами замкнуты на окружности: $\Delta x_k = x_{k+1} - x_k$, $k = 1, 2, \ldots, N-1$ и $\Delta x_N = 1 - x_N + x_1$ (Бирюков и Теребиж, 1989). Структура $E(\nu)$ связана с определением (2.5.1) энтропии дискретной случайной величины. Предположим, что нужно оценить степень равномерности заданного точечного распределения $\{x_k\}$. Введем для этого случайную величину ξ - номер того интервала из совокупности $\{\Delta x_k\}$, на который попадет точка при случайном равномерном ее бросании на отрезок $[0,1)$. Очевидно, ξ принимает значения $k = 1, 2, \ldots, N$ с вероятностями $\{p_k = \Delta x_k\}$, так что ее энтропия равна $S(\xi) = -\sum_{1}^{N} \Delta x_k \log \Delta x_k$. Максимального значения $S_{max} = \log N$ энтропия достигает при строго равномерном распределении $\Delta x_k = 1/N$. Цель принятой в (6) нормировки заключается в том, чтобы значения $E(\nu)$ находились в интервале от 0 до 1, причем существенные отклонения от равномерности $\{x_k\}$ соответствовали максимумам периодограммы $E(\nu)$. При помощи методов адресной сортировки, которые позволяют провести упорядочение фаз за время, пропорциональное N, можно вычислить $E(\nu)$ для достаточно длинных импульсных последовательностей, однако главную область приложений этой статистики составляют короткие ряды, где биновое разбиение оказывается слишком грубым. К таким рядам относятся, в частности, последовательности импульсов в детекторах гравитационных волн и нейтрино.

Особенностью изложенного подхода является нечувствительность статистик $\chi^2(\nu)$ и $E(\nu)$ к упорядоченности *последовательных* значений $\Delta N_i \equiv N_i - N\Delta x_i$ или Δx_k. Действительно, величина $\chi^2(\nu)$ не изменится, если бины произвольно переставить, сохранив наблюдаемые отклонения ΔN_i в каждом из них.

Между тем, именно наличие порядка в вариациях $\{\Delta N_i\}$ или $\{\Delta x_k\}$ могло бы служить указанием на реальность периодичности (скажем, при глазомерном оценивании фазовых диаграмм). Хорошей иллюстрацией к сказанному является пуассоновская последовательность со ступенчатым изменением средней интенсивности (пример 5.1.2, § 5.4). Учесть упорядоченность такого рода несложно *при наличии предварительной информации* об изучаемом процессе. Так, если известна приблизительная форма кривой блеска, то описанные критерии можно дополнить статистикой, характеризующей отклонение от ожидаемой последовательности событий. Фактически так и происходит при построении импульсного спектра $I_N(\nu)$, где пробной кривой блеска является синусоида.

Как и ранее, периодограмма Фурье - импульсный спектр - и периодограммы непараметрических методов отличаются в том отношении, что первая из них не обнаруживает дополнительных экстремумов на частотах субгармоник $\nu_0/2$, $\nu_0/3$, $\nu_0/4$,... основного периода (см. § 4.5).

Выделение характерных масштабов переменности. Как импульсный спектр, так и периодограммы $\chi^2(\nu)$ и $E(\nu)$ предназначены главным образом для поиска близких к периодическим компонентов в случайной последовательности импульсов. Допустим однако, что мы имеем дело с такими колебаниями средней интенсивности процесса, которые не сохраняют с течением времени амплитуду и фазу. В частности, это могут быть узкополосные случайные колебания с характерным временным масштабом $\tau_n \equiv \nu_n^{-1}$ (§ 5.5). На фазовой диаграмме такие колебания размываются ввиду набега фазы далеко отстоящих участков ряда. Для выделения стохастической переменности такого рода применяется следующий подход.

Разделим весь промежуток наблюдения T на R интервалов продолжительности $\tau = T/R = \nu^{-1}$ и образуем величину

$$\chi^2(\tau) = \sum_{j=1}^{R} \frac{[\mathfrak{N}_j(\tau) - M(\tau)]^2}{M(\tau)} . \qquad (5.3.7)$$

Можно ожидать, что $\chi^2(\tau)$ по-разному ведет себя в областях $\tau \leqslant \tau_n$ и $\tau \geqslant \tau_n$, и это обстоятельство позволит оценить значение параметра τ_n. Мы видим, что здесь также строится пе-

риодограмма χ^2_ν, но на бины разбивается весь интервал наблюдений, а количество бинов однозначно связано с пробной частотой: $R(\nu) = \nu T = N/M(\tau)$ (мы отвлекаемся от легко учитываемых краевых эффектов).

Описанный подход неоднократно использовался при анализе результатов физических экспериментов (см. Иэди и др., 1971), а в астрономии, по-видимому,- начиная с работы Болдта и др. (1971). Сведения по применениям к счету фотонов в рентгеновской области содержатся в статьях Формэна и др. (1976), Коллуры и др. (1987), в оптическом диапазоне - статьях Шварцмана и его сотрудников (Плахотниченко, 1983; Шварцман и др., 1988).

Для того чтобы выяснить связь величины (7) с традиционными статистиками дисперсионного анализа, усредним $\chi^2(\tau)$ по всем значениям $\mathfrak{N}_j$. Вспоминая определение (5.1.11) индекса дисперсии, получаем

$$\frac{1}{R}\,\langle\chi^2(\tau)\rangle = v_{\mathfrak{N}}(\tau). \qquad (5.3.8)$$

Таким образом, о ж и д а е м о е о т н о ш е н и е χ^2 к ч и с л у б и н о в р а в н о и н д е к с у д и с п е р с и и . Явное представление последнего через автоковариационную функцию $\psi_n(\tau)$ и спектральную плотность $g_n(\nu)$ колебаний средней интенсивности процесса $n(t)$ для весьма широкого класса импульсных последовательностей - дважды стохастического процесса Пуассона - дается формулами (5.1.41) и (5.1.42). Ввиду (8) далее для определенности говорится о $v_{\mathfrak{N}}(\tau)$.

Заметим, что перед вычислениями дисперсии следует устранить общий тренд временного ряда. Далее, при низком отношении сигнала к шуму на периодограмме сказывается выбор подразделения интервала наблюдений на бины. В этой связи Коллура и др. (1987) рекомендуют провести усреднение по начальной фазе биновой структуры. Наконец, вводя эффективное время наблюдений в каждом бине, нетрудно модифицировать (7) так, чтобы учесть пропуски экспозиции.

Как было показано в § 5.1, сопоставление наблюдемой зависимости $v_{\mathfrak{N}}(\tau)$ с ожидаемым поведением индекса дисперсии (5.1.47) позволяет приближенно оценить характерное время

переменности τ_n. Более точная оценка τ_n может быть получена сравнением с результатами модельных расчетов $v_{\mathfrak{N}}(\tau)$ по формулам (5.1.41) или (5.1.42).

После нахождения τ_n, казалось бы, по тем же точным или асимптотическим формулам для $v_{\mathfrak{N}}(\tau)$ можно оценить и амплитуду переменности σ_n. Однако такая оценка сама по себе мало говорит о физических свойствах переменного источника, поскольку она зависит от наблюдаемого среднего уровня интенсивности $\bar{n}$. Таким образом, вместо $v_{\mathfrak{N}}(\tau)$ желательно оперировать с функцией, зависящей только от абсолютных, собственных свойств источника. Пимонов и Теребиж (1981) показали, что существует множество таких функций $H_1, H_2, \ldots$, и выразили их через моменты $\mathfrak{N}$ (см. § 5.5). Для пуассоновской последовательности с постоянной средней интенсивностью все эти функции, названные в указанной работе *инвариантами фильтра*, тождественно равны 1. В общем же случае лишь $H_1(\tau) \equiv 1$, а дальнейшие инварианты определяются формулой (5.5.30). В частности, второй инвариант $H_2(\tau)$ равен

$$H_2(\tau) = 1 + \frac{v_{\mathfrak{N}}(\tau) - 1}{M(\tau)} . \qquad (5.3.9)$$

Подставляя сюда выражения (5.1.41) и (5.1.42), получаем:

$$H_2(\tau) = 1 + \frac{2}{(\bar{n}\tau)^2} \int_0^\tau (\tau - \tau')\psi_n(\tau')\,d\tau' =$$

$$= 1 + (\bar{n})^{-2} \int_{-\infty}^{\infty} \operatorname{sinc}^2(\nu\tau)\, g_n(\nu)\,d\nu. \qquad (5.3.10)$$

Асимптотическое поведение $H_2(\tau)$ определяется формулами (9) и (5.1.47):

$$H_2(\tau) \simeq \begin{cases} 1 + (\sigma_n/\bar{n})^2, & \tau \ll \tau_n, \\ 1 + \dfrac{2B}{(\bar{n})^2} \dfrac{1}{\tau}\left(1 - \dfrac{\tau_n}{\tau}\right), & \tau \gg \tau_n. \end{cases} \qquad (5.3.11)$$

Функция $H_2(\tau)$ действительно содержит только не зависящие от условий наблюдения величины типа относительной амплитуды переменности $\sigma_n/\bar{n}$ и характерного временного масштаба τ_n. Более тонкие свойства процесса определяются инвариантами высоких порядков.

Заметим, что введенная Бернштейном и Шварцманом (1983) d_2-функция (фактическое использование d_2-функции в эксперименте МАНИЯ началось за несколько лет до этого) и введенный Пимоновым и Теребижем (1981) второй инвариант $H_2(\tau)$ отличаются от статистики χ^2 центрированием к единице и нормированием на среднее значение. Формула (10) была получена в цитированной статье Пимонова и Теребижа, где найдена также дисперсия выборочного значения $\hat{H}_2$, необходимая для вычисления значимости наблюдаемых отклонений:

$$D(\hat{H}_2) \simeq \frac{2}{(R-1)\,(\bar{n})^2}. \qquad (5.3.12)$$

Модельные расчеты $H_2(\tau)$ описаны в следующем параграфе.

Характеризуя в общем дисперсионные методы выделения стохастической переменности, нужно сказать, что ожидать от них высокой эффективности не приходится. Причина этого видна из соотношения (10): необходимая оценка спектральной плотности колебаний $n(t)$ *интегральным* образом связана с дисперсией импульсного процесса. Поэтому восстановление $g_n(\nu)$ является трудной, к тому же некорректной задачей (см. § 3.8). По-видимому, более эффективны прямые оценки $g_n(\nu)$, основанные на построении импульсного спектра.

§ 5.4 Примеры

Отличительная черта развитого выше общего подхода состоит в том, что усреднение по случайному распределению событий - пуассоновскому шуму - проводится на раннем этапе исследования. После этого основные уравнения включают в себя только статистические характеристики колебаний *средней* интенсивности $n(t)$. Это позволяет сосредоточить внимание на вопросах, которые, собственно, и представляют интерес для наблюдателя.

Иллюстрируем введенные понятия несколькими примерами, когда $n(t)$ изменяется заданным образом - детерминированно или стохастически. Полученные соотношения полезны как сами по себе, так и в связи с тем, что они позволяют качественно интерпретировать переменность более сложного характера.

Гармонический пуассоновский процесс был определен в

§ 5.1 соотношением

$$n(t) = \bar{n}[1+\mu\cos(2\pi\nu_0 t-\varphi_0)], \qquad (5.4.1)$$

где коэффициент модуляции μ предполагается по модулю меньшим 1.

Найдем сначала импульсный спектр гармонического процесса Пуассона конечной длительности T. Согласно (1.4.22) спектральная плотность переменного слагаемого $n(t)$ равна

$$g_n^{(T)}(\nu) \simeq T\,\frac{(\bar{n}\mu)^2}{4}\,\mathrm{sinc}^2\Big[(\nu-\nu_0)T\Big], \quad \nu \geqslant 0, \qquad (5.4.2)$$

где для упрощения записи мы ограничились положительной областью частот и считаем ряд достаточно длинным ($T^{-1} \ll \nu_0$). Подставляя это выражение в (5.2.23) и учитывая, что полное количество зарегистрированных за время T импульсов $N = \bar{n}T$, получаем для импульсного спектра, обусловленного модуляцией, следующее выражение:

$$\Delta I_N(\nu) \simeq N\,\frac{\mu^2}{4}\,\mathrm{sinc}^2\Big[(\nu-\nu_0)T\Big], \quad \nu \geqslant 0. \qquad (5.4.3)$$

При $T \gg P_0 \equiv \nu_0^{-1}$ добавочный спектр состоит из двух симметрично расположенных относительно точки $\nu = 0$ линий с шириной $\sim T^{-1}$ и амплитудой

$$\Delta I_N(\nu_0) \simeq N\,\frac{\mu^2}{4}\,. \qquad (5.4.4)$$

Последнее соотношение позволяет конкретизировать условия, необходимые для надежного выделения периодичности в импульсной последовательности. Именно, при обнаружении с уровнем значимости q должно выполняться неравенство

$$N\,\frac{\mu^2}{4} > I_q - 1, \qquad (5.4.5)$$

где I_q - критическое значение импульсного спектра, зависящее от q, спектрального разрешения и ширины исследуемого интервала частот (см. §§ 5.2 и 2.3). Условие (5) иллюстрируется рис. 5.4.1; например, для выделения 10 %-ной модуляции при $I_q = 11$ нужно накопить не менее 4 000 импульсов.

Предположим теперь, что та же случайная последовательность импульсов анализируется при помощи метода χ^2. Найдем приращение $\Delta\chi^2(\nu_0)$ этой статистики вблизи основной частоты

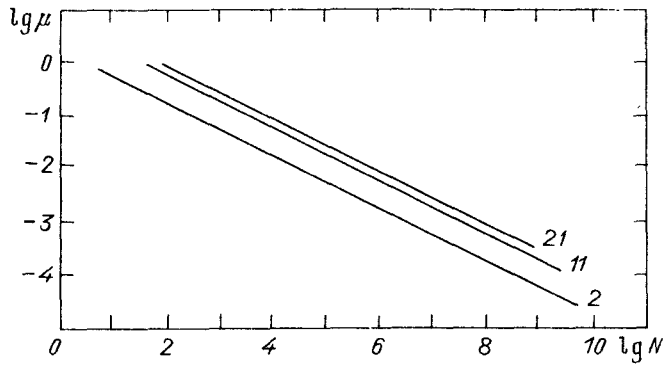


Рис. 5.4.1. Критические области для выделения гармонического компонента в случайной последовательности импульсов в зависимости от количества отсчетов N и глубины модуляции μ

колебаний. При сворачивании ряда со средней интенсивностью (1) и пробным периодом $P_0 = \nu_0^{-1}$ интервал фаз Δx_i содержит приблизительно $N_i = N\Delta x_i[1 + \mu\cos(2\pi x_i)]$ отсчетов. Ожидаемое при равномерном распределении количество отсчетов в Δx_i равно $N\Delta x_i$, так что отклонение $N_i - N\Delta x_i \simeq N\mu\Delta x_i\cos(2\pi x_i)$. Полагая здесь для простоты все фазовые интервалы одинаковыми, т.е. $\Delta x_i \equiv r^{-1}$, $x_i = i/r$, подставляя эти выражения в (5.3.4) и принимая во внимание (П.1.5), находим:

$$\Delta\chi^2(\nu_0) \simeq N\frac{\mu^2}{2}. \qquad (5.4.6)$$

Более точные выкладки Лихи и др. (1983, *a*), Лихи (1987) показывают, что это выражение содержит еще поправочный множитель $c_r = \mathrm{sinc}^2(r-1)$, зависящий от количества фазовых интервалов. Поскольку при $x \ll 1$ функция $\mathrm{sinc}(x) \simeq 1-(\pi x)^2/6$, фактор c_r следует учитывать лишь при небольшом количестве бинов (уже $c_{10} \approx 0{,}97$). Форма спектральной линии $\Delta\chi^2(\nu)$ определяется той же функцией $\mathrm{sinc}^2 x$, которая входит в (3); ограничиваясь вновь областью $\nu \geqslant 0$, запишем вклад гармонического компонента следующим образом:

$$\Delta\chi^2(\nu) \simeq N\frac{\mu^2}{2}c_r\mathrm{sinc}^2\left[(\nu-\nu_0)T\right]. \qquad (5.4.7)$$

Формулы (3) и (7) во многом сходны. Следует однако иметь

в виду, что для хаотического фазового распределения $\langle I_N \rangle = 1$, в то время как $\langle \chi^2 \rangle = r-1$. Поэтому для выделения периодичности необходимо, чтобы отношение $N\mu^2/2(r-1)$ было велико. Условие, аналогичное (5), имеет вид

$$N \frac{\mu^2}{2} c_r > \chi_q^2 - (r-1), \qquad (5.4.8)$$

где χ_q^2 - критическое значение случайной величины, подчиняющейся χ^2-распределению с $r-1$ степенями свободы.

В первом приближении оценки периода и коэффициента модуляции можно найти, исходя из положения и амплитуды максимума периодограммы, т.е. на основании соотношений (4) или (6). Дальнейшее уточнение оценок производится путем подгонки теоретической формы линии $\mathrm{sinc}^2(\nu-\nu_0)T$ к наблюдаемому отклику соответствующей периодограммы (Лихи, 1987).

Полезно рассмотреть численный пример случайной импульсной последовательности. Гармонический пуассоновский процесс был реализован при помощи описанного в § 5.1 метода фон Неймана. Именно, алгоритм вычислений состоял из следующих шагов: 1) задаются средняя интенсивность процесса $\bar{n}$, коэффициент модуляции μ, угловая частота $\omega_0 = 2\pi\nu_0$ и начальная точка отсчета времени t (начальная фаза $\varphi_0 = 0$); 2) генерируются две случайные величины α_1 и α_2, равномерно распределенные на $[0,1]$; 3) вычисляется пробный интервал времени до следующего события $\underline{X} = -\ln\alpha_1/\bar{n}(1-\mu)$; 4) вычисляется $\eta = \bar{n}(1+\mu)\alpha_1\alpha_2$; 5) вычисляется $f = Ae^{-B}$, где $A = \bar{n}(1+\mu\cos(\omega_0(t+\underline{X})))$, $B = \bar{n}(\underline{X}+\mu\omega_0^{-1}(\sin(\omega_0(t+\underline{X})) - \sin(\omega_0 t)))$; 6) если $\eta > f$, то выполняется переход к пункту 2, если $\eta \leqslant f$, то принимается $\underline{X}$ в качестве свободного интервала до следующего события, t заменяется на $t+\underline{X}$, выполняется переход к пункту 2; 7) цикл заканчивается по достижении заданного количества N отсчетов $\{t_k\}$.

По формулам (5.2.24) и (5.3.4) были вычислены импульсный спектр $I_N(\nu)$ и периодограмма $\chi^2(\nu)$ одной и той же реализации гармонического процесса Пуассона при следующих значениях параметров: $\bar{n} = 10\ \mathrm{c}^{-1}$, $\mu = 1/4$, $P_0 = e \approx 2,71828$ с, $t_0 = 0$, $N = 1000$, $r = 10$. Таким образом, свободный интервал между импульсами в среднем составлял 0,1 с, полное время наблюдений $T \sim 100$ с, среднее количество событий на пе-

риод $\gamma \equiv \bar{n}P_0 \approx 27,2$. Ожидаемые согласно (4) и (6) при $\nu_0 \approx$ $\approx 0,368\ c^{-1}$ значения $\Delta I_N(\nu_0)$ и $\Delta\chi^2(\nu_0)$ равны соответственно 16 и 31. Результаты расчетов представлены на рис. 5.4.2. Наблюдаемые для реализованной последовательности пики несколько больше возвышаются над уровнем шума. Отношение сигнала к шуму составляет ≈20 и ≈4 для импульсного спектра и χ^2 соответственно. Следует обратить внимание на появление максимумов $\chi^2(\nu)$ вблизи частот субгармоник $\nu_0/2$, $\nu_0/3, \ldots$, о которых говорилось в §§ 5.3 и 4.5.

Для проверки чувствительности импульсного спектра к высокочастотным колебаниям расчеты были проведены также для $P_0 = 0,075$ с и прежних значений остальных параметров. В этом случае среднее количество событий на период γ составляет лишь 3/4. Отклик на частоте $\nu_0 \approx 13,333\ c^{-1}$ оказался равным $I_N(\nu_0) \approx 22,5$ при сохранении прежнего уровня шума (см. рис. 5.4.2), т.е. импульсный спектр столь же эффективен при выделении быстрой периодичности, как и при обнаружении плавных колебаний средней интенсивности.

При обоих значениях периода были построены гистограммы длин интервалов между событиями ($r = 23$). Наблюдаемые распределения длин свободных интервалов с высокой надежностью оказались неотличимыми от экспоненциального распределения, присущего пуассоновскому потоку постоянной интенсивности (заметим в этой связи, что сама по себе даже строгая экспоненциальность $f_X(\tau)$ не свидетельствует о том, что поток пуассоновский с постоянной интенсивностью - необходима еще независимость последовательных интервалов; примером может служить простейший пуассоновский поток со свободными интервалами, разделенными точно пополам).

Для двух описанных выше, а также для ряда других примеров импульсных последовательностей с относительно невысоким отношением сигнала к шуму мы сопоставили эффективность методов, основанных на вычислении $I_N(\nu)$ и периодограммы свободных интервалов $P_X(z)$, вычисленной согласно (5.2.16). Во всех случаях $P_X(z)$ не обнаружила свидетельств периодических колебаний $n(t)$, в то время как отклик импульсного спектра был вполне определенным.

Таким образом, в условиях описанных выше примеров среди

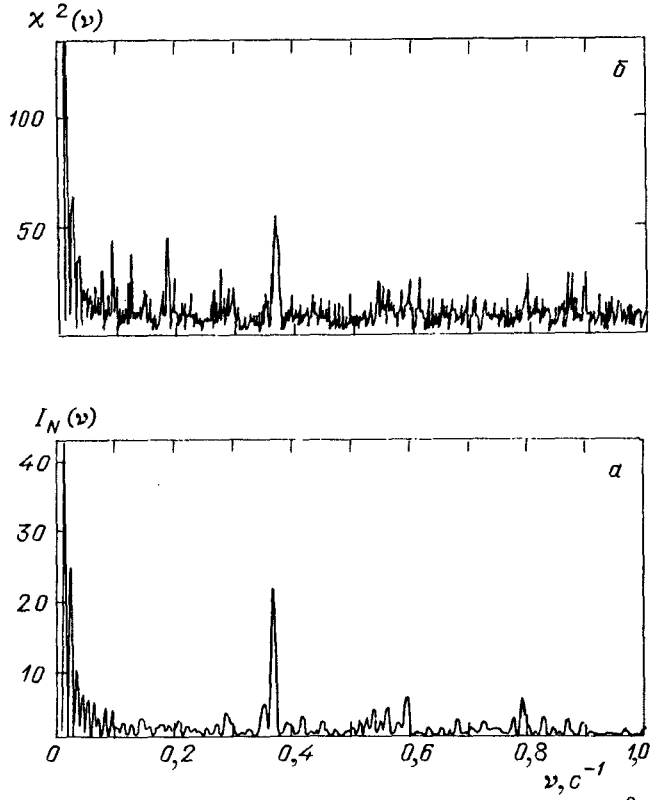


Рис. 5.4.2. Импульсный спектр (а) и периодограмма χ^2 (б) одной и той же случайной последовательности отсчетов, содержащей периодический компонент с частотой $\nu_0 \simeq 0{,}368$ Гц.

всех перечисленных методов безоговорочное предпочтение следует отдать расчету импульсного спектра.

Ступенчатый пуассоновский процесс был определен в примере 5.1.2, поведение его средней интенсивности $n(t)$ представлено на рис. 5.1.3. Поскольку спектр ступенчатого процесса имеет максимумы и на частотах гармоник основного периода, рассмотрение этого процесса весьма полезно для выяснения общих закономерностей, присущих анализу последовательностей импульсов.

Заметим прежде всего, что из (1.6.7) и (5.2.23) следует простое соотношение между средним приращением импульсного спектра и периодограммой вариации $n(t)$:

$$\Delta I_N(\nu) \simeq \frac{T^2}{N} P_T(\nu). \qquad (5.4.9)$$

Учитывая (1.6.9), находим для частот гармоник

$$\Delta I_N(\nu_k) \simeq \frac{T^2}{N} \frac{2a_k^2 + b_k^2}{4}, \qquad k = 1, 2, \ldots, \qquad (5.4.10)$$

где a_k и b_k - коэффициенты Фурье $n(t)$. Для ступенчатой функции эти коэффициенты были найдены в примере 1.4.1; согласно (1.4.33),

$$a_k = 2(n_2 - n_1)\delta \operatorname{sinc}(k\delta), \quad b_k = 0. \qquad (5.4.11)$$

Подставляя эти значения в (10) и обозначая посредством $\alpha \equiv (n_2 - n_1)/n_1$ относительную амплитуду вспышки, получаем окончательно:

$$\Delta I_N(k\nu_0) = N\left[\frac{\alpha\delta}{1+\alpha\delta} \operatorname{sinc}(k\delta)\right]^2, \qquad k = 1, 2, \ldots \qquad (5.4.12)$$

В случае $k = 1$ соотношение, эквивалентное (12), приводят Лихи и др. (1983, *b*).

Как и должно быть, при $\alpha = 0$ или крайних значениях относительной продолжительности вспышки ($\delta = 0$; 1), когда средняя интенсивность процесса становится постоянной, добавочный спектр исчезает. Формула (12) позволяет оценить коэффициент скважности δ по отношению амплитуд I_N на частотах гармоник, а сами амплитуды $I_N(k\nu_0)$ дают и значение α. Конечно, для такого анализа желательно, чтобы отношение сигнала к шуму было достаточно высоким, т.е. последовательность содержала много отсчетов. Если это условие выполнено, и тем не менее высокие гармоники не обнаружены, следует полагать, что вариации средней интенсивности лучше описываются гармоническим законом (1).

Приведем без вывода формулу для приращения χ^2-периодограммы на основной частоте ν_0 в случае ступенчатой модуляции средней интенсивности (Лихи и др., 1983, *b*):

$$\Delta\chi^2(\nu_0) = N\left(\frac{\alpha\delta}{1+\alpha\delta}\right)^2 \times \begin{cases} r - 1 - \delta r^2/3, & 0 \leqslant \delta \leqslant r^{-1}, \\ (1 - \delta - r\delta/3)/\delta, & r^{-1} \leqslant \delta \leqslant 1 - r^{-1}. \end{cases} \qquad (5.4.13)$$

Для обоих частных примеров изменения $n(t)$ отношения

$\Delta I_N(\nu_0)/N$ и $\Delta\chi^2(\nu_0)/N$ пропорциональны мощности относительной вариации средней интенсивности. Этот вывод справедлив для произвольного закона изменения $n(t)$.

Стохастическая переменность. Обратимся к вопросу о выделении случайных колебаний средней интенсивности процесса $n(t)$. Запишем основное соотношение (5.3.10) в виде

$$H_2(\tau) = 1+ph_2(\tau), \quad p = (\sigma_n/\bar{n})^2,$$

$$h_2(\tau) = \frac{2}{(\sigma_n\tau)^2}\int_0^\tau(\tau-\tau')\psi_n(\tau')d\tau'. \qquad (5.4.14)$$

В соответствии с (5.3.11) имеем $h_2(0) = 1$, $h_2(\infty) = 0$, так что параметр p характеризует *относительную мощность* переменности. Найдем p и $h_2(\tau)$ для некоторых конкретных моделей переменности $n(t)$.

Модель А. Предположим, что $n(t)$ может принимать только два значения: n_1 (основное состояние) и $n_2 \geqslant n_1$ (возбужденное состояние, вспышка). Состояния последовательно сменяют друг друга, продолжительности каждого из них - независимые случайные величины, подчиняющиеся экспоненциальному распределению (П.4.19) со средними значениями τ_1 и τ_2 для основного и возбужденного уровней соответственно. Таким образом, средняя доля времени, в течение которого источник находится в возбужденном состоянии, равна $\delta = \tau_2/(\tau_1+\tau_2)$. Средняя интенсивность процесса $\bar{n} = (1-\delta)n_1+\delta n_2$. За неимением места опустим вывод автоковариационной функции процесса; окончательный результат имеет вид

$$\psi_n(\tau) = (n_2-n_1)^2\delta(1-\delta)e^{-|\tau|/\tau_n}, \qquad (5.4.15)$$

$$\tau_n = (\tau_1^{-1}+\tau_2^{-1})^{-1}.$$

Мы видим, что характерное время переменности τ_n представляет собой среднее гармоническое величин τ_1 и τ_2, а дисперсия процесса $\sigma_n^2 = \psi_n(0) = (n_2-n_1)^2\delta(1-\delta)$. Экспоненциальная форма $\psi_n(\tau)$ нами уже рассматривалась в примере 5.1.4. Согласно (5.1.44), (5.3.9) и (14) имеем

$$p = \frac{\alpha^2\delta(1-\delta)}{(1+\alpha\delta)^2}\ , \quad \alpha = \frac{n_2-n_1}{n_1}\ ,$$
$$h_2(\tau) = 2u^{-2}(e^{-u}-1+u), \quad u = \tau/\tau_n \qquad (5.4.16)$$

(Пимонов и Теребиж, 1981). Функция $h_2(\tau)$ представлена на рис. 5.4.3, ее асимптотики имеют вид: $h_2(\tau) \simeq 1-u/3$ при $u \to 0$ и $h_2(\tau) \simeq 2u^{-1}$ при $u \to \infty$.

Из (5.1.45) и (5.2.23) следует, что нормированный к средней интенсивности процесса импульсный спектр в данном случае равен

$$(\bar{n})^{-1} I_N(\nu) = \frac{2p\tau_n}{1+(2\pi\tau_n\nu)^2}. \qquad (5.4.17)$$

Величина $(\bar{n})^{-1} I_N(\nu)$ является инвариантом подобно $H_2(\tau)$. Соотношение (17) в той же мере можно использовать для нахождения характеристик переменности, как и (16).

Модель В. Пусть $n(t)$ представляет собой сумму постоянной величины n_0 - интенсивности основного состояния - и слагаемого, отвечающего «вспышкам» детерминированной формы $s(u)$. Вспышки происходят в случайные моменты времени $\{t_k\}$, образующие пуассоновскую последовательность; среднюю частоту вспышек обозначим через n_*. Амплитуды $\alpha_k = \Delta n_k/n_0$ вспышек предполагаются случайными с произвольной плотностью распределения. Характерная длительность вспышек τ_n считается для простоты фиксированной величиной (вспышки могут произвольным образом накладываться друг на друга).

Формальное описание процесса дается соотношением

$$n(t) = n_0\left[1+\sum_k \alpha_k s\left(\frac{t-t_k}{\tau_n}\right)\right], \qquad (5.4.18)$$

где без потери общности можно считать $\int_{-\infty}^{\infty} s(u)du = 1$. Мы видим, что модель В соответствует обсуждавшемуся в § 2.8 дробовому шуму. Согласно формулам Кемпбелла (2.8.5) мы имеем:

$$\bar{n} = n_0(1+n_*\bar{\alpha}\tau_n),$$
$$\psi_n(\tau) = n_0^2 n_*\overline{\alpha^2}\tau_n\Phi(u), \quad u = \tau/\tau_n, \qquad (5.4.19)$$

$$\Psi(u) = \int_{-\infty}^{\infty} s(u')s(u+u')du'.$$

Из (14) и (19) находим для модели *B*

$$p = \frac{n_*\overline{\alpha^2}\tau_n\Psi(0)}{(1+n_*\bar{\alpha}\tau_n)^2}, \qquad (5.4.20)$$

$$h_2(\tau) = \frac{2}{\Psi(0)u^2}\int_0^u (u-u')\Psi(u')du'.$$

Если учесть, что $n_*\tau_n$ представляет собой в данном случае долю времени, проводимого источником в возбужденном состоянии, то между (20) и (16) нетрудно усмотреть сходство во многих отношениях. Для того чтобы более полно сопоставить модели *A* и *B*, конкретизируем форму вспышек $s(u)$ в (20).

Пусть сначала $s(u) = 1$ при $|u| \leqslant 1/2$, $s(u) = 0$ при $|u| > 1/2$, т.е. вспышки имеют прямоугольную форму. В примере 1.3.2 было найдено, что при этом $\Psi(u) = 1-|u|$ для $|u| \leqslant 1$ и $\Psi(u) = 0$ для $|u| > 1$. Наконец, второе из соотношений (20) принимает вид

$$h_2(\tau) = \begin{cases} 1 - \frac{u}{3}, & u \leqslant 1, \\ \frac{1}{u}\left(1 - \frac{1}{3u}\right), & u > 1. \end{cases} \qquad (5.4.21)$$

Эта функция также представлена на рис. 5.4.3; обратим внимание, что точная формула (21) близка к асимптотическому представлению $h_2(\tau)$ в модели *A*.

Для оценки степени зависимости результатов от формы вспышек рассмотрим еще $s(u)$ вида гауссовой плотности (1.3.16) при $\sigma^2 = 1$. Функция $\Psi(u)$ также имеет вид плотности Гаусса (1.3.18) при $\sigma_0^2 = 2$. В данном случае $\Psi(0) = (2\sqrt{\pi})^{-1}$ и

$$h_2(\tau) = 4u^{-1}\left[\int_0^{u/2} e^{-x^2}dx - u^{-1}(1-e^{-u^2/4})\right] \qquad (5.4.22)$$

(рис. 5.4.3). Напомним, что здесь, как и всюду в этом параграфе, $u = \tau/\tau_n$.

Инвариантный импульсный спектр для модели *B* описывается

выражение.

$$I_N(\nu)/\bar{n} = \frac{n_* \overline{a^2} \tau_n^2}{(1+n_* \bar{a} \tau_n)^2} \, |S(\nu\tau_n)|^2 , \qquad (5.4.23)$$

где $S(z)$ - образ Фурье функции $s(u)$. В частности, при гауссовой форме вспышки

$$I_N(\nu)/\bar{n} = 2\sqrt{\pi}\, p\tau_n e^{-(2\pi\tau_n\nu)^2} . \qquad (5.4.24)$$

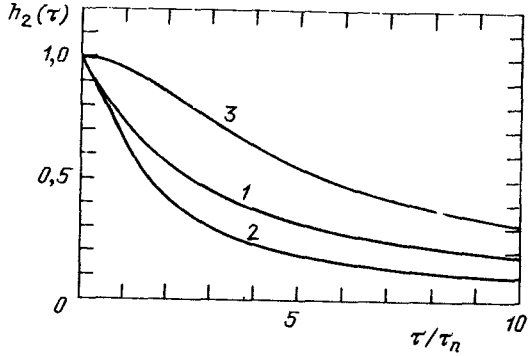


Рис. 5.4.3. Вид функции $h_2(\tau)$ для модели A(1), а также вспышек прямоугольной (2) и гауссовой (3) формы в модели B

Как и ранее, ширина наблюдаемого спектра и его центральная интенсивность позволяют оценить мощность и характерный временной масштаб переменности источника.

§ 5.5. Счет фотонов

Уже три десятилетия счет отдельных фотонов является основным методом изучения слабых астрономических объектов в высокочастотной области спектра, включающей и оптический диапазон.

Если отвлечься от второстепенных деталей, то всякий приемник излучения преобразует входной поток фотонов $n_\phi(c^{-1})$ в некоторый сигнал $v(t)$ на выходе (рис. 5.5.1). Мы рассматриваем стационарное поле излучения, так что средние значения $\bar{n}_\phi$ и $\bar{v}$ не зависят от времени, а естественные как при квантовом, так и при классическом описании флуктуации n_ϕ приводят к флуктуациям выходного сигнала $v(t)$.

Качество приемника определяется следующими основными

факторами: 1) средним значением $\bar{v}$ при данном $\bar{n}_\phi$, освобожденном от темнового потока $\bar{v}_d$ (желательно, чтобы как можно большая часть фотонов вносила вклад в регистрируемый сигнал, т.е. *квантовый выход*, пропорциональный $\bar{v}/\bar{n}_\phi$, был как можно более высоким); 2) шириной того участка значений входного потока, в пределах которого можно считать $\bar{v}$ пропорциональным $\bar{n}_\phi$, т.е. *динамическим диапазоном* приемника; 3) дисперсией σ_v^2 выходного сигнала; если она мала, то полагают, что световой поток измеряется с высокой точностью (заметим, впрочем, что этот общепринятый подход имеет ограниченную область применения; корректная трактовка дана в § 5.6).

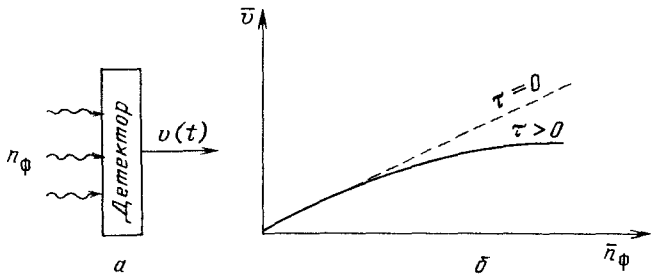


Рис. 5.5.1. Схема детектирования фотонов (а) и зависимость среднего сигнала $\bar{v}$ на выходе детектора от средней интенсивности падающего потока $\bar{n}_\phi$ (б)

Конечной задачей наблюдателя является получение *оценки* $\hat{n}_\phi$ истинного значения среднего потока $\bar{n}_\phi$, и перечисленные факторы определяют точность $\hat{n}_\phi$ при фиксированном времени наблюдения T.

Почему же счет фотонов стал основным методом исследования слабых световых потоков? Конечно, есть преимущества типа непосредственной цифровой регистрации отсчетов, однако вполне возможна и реализация в цифровом виде аналогового сигнала, например, тока на выходе фотоумножителя (ФЭУ). Существо же дела заключается в том, что дисперсия выходного сигнала при счете фотонов принципиально меньше, чем при любом другом способе регистрации, - она ограничена только

неизбежными квантовыми флуктуациями потока фотонов.

Приведем два примера. При счете фотонов с помощью ФЭУ выходным сигналом является близкий к пуассоновскому поток стандартизованных по форме и амплитуде импульсов напряжения со средней скоростью счета $\bar{n}_{имп}(с^{-1})$. Дисперсия накопленного за время T количества импульсов примерно равна $\bar{n}_{имп}\,T$. Если же регистрируется аналоговый сигнал - ток на выходе ФЭУ, - то он включает в себя не только пропорциональное интенсивности света количество импульсов (как в первом случае), но и сумму амплитуд этих импульсов. Амплитудное распределение импульсов ФЭУ является очень широким, поэтому относительная дисперсия тока велика и аналоговая регистрация менее точна (приблизительно в 1,5 раза), чем счет событий. Другой пример касается таких современных приемников, *как приборы с зарядовой связью* (ПЗС). Основные достоинства ПЗС заключаются в возможности реализовать многоканальную регистрацию и в высоком квантовом выходе. Вместе с тем, наличие шума считывания при работе с ПЗС чаще всего не позволяет каждому элементу достигнуть эффективности, при которой отношение сигнала к шуму ограничено лишь фотонной статистикой. Поэтому счетчики фотонов продолжают разрабатываться (см., например, обзор Тимоти, 1988) и применяться в условиях, когда требуется предельно высокая квантовая эффективность обнаружения (§ 5.6).

В связи с приведенными примерами заметим, что многие наблюдатели знают и постоянно используют тот факт, что статистика фотоотсчетов близка к пуассоновской. Но причина этого явления и условия, при которых оно осуществляется (даже при «идеальном» эксперименте), известны совершенно недостаточно. Между тем вопрос нетривиален, так как сами по себе фотоны подчиняются *статистике Бозе-Эйнштейна*. С другой стороны, вопрос имеет не только академический интерес, поскольку, например, именно в рамках теории счета фотонов ясно интерпретируется известный эксперимент Брауна и Твисса (1958) по измерению звездных диаметров при помощи корреляции интенсивности.

Помимо астрономии техника счета фотонов получила широкое

распространение в спектроскопии высокого разрешения, исследованиях лазерного излучения и других областях физики. Рассматриваемую область невозможно осветить сколько-нибудь полно в пределах одного параграфа. Мы ограничимся здесь обсуждением принципиальных вопросов и примерами. Изложение основано на классическом приближении теории счета фотонов, которое дает правильное количественное описание важнейших экспериментов. Более полное изложение можно найти в обзоре Меты (1970), монографии Лоудона (1973) и сборнике под редакцией Камминса и Пайка (1974).

Спектр и автоковариационная функция электромагнитных колебаний. Предположим, что стационарный световой поток создается вблизи детектора некогерентным источником излучения. Будем рассматривать вещественную функцию $\xi(t)$ как скалярный аналог напряженности электрического поля в фиксированной точке пространства. Поскольку мы обычно имеем дело с суперпозицией переменных полей очень большого количества атомов, можно трактовать $\xi(t)$ как стационарный случайный процесс с нулевым средним и автоковариационной функций

$$\psi_\xi(\tau) = \langle \xi(t)\xi(t+\tau)\rangle = \int_{-\infty}^{\infty} e^{i\cdot 2\pi\nu\tau} g_\xi(\nu)\,d\nu, \qquad (5.5.1)$$

где $g_\xi(\nu) = g_\xi(-\nu)$ - *спектральная плотность электромагнитных колебаний* (см. § 1.5). Обращение формулы (1), выражающей теорему Винера Хинчина, дает:

$$g_\xi(\nu) = \int_{-\infty}^{\infty} e^{-i\cdot 2\pi\nu\tau}\psi_\xi(\tau)\,d\tau. \qquad (5.5.2)$$

Определим далее *мгновенную интенсивность излучения* $J(t)$ квадратичной функцией напряженности поля:

$$J(t) = \frac{1}{2}\,\xi^2(t). \qquad (5.5.3)$$

Очевидно, $J(t)$ также представляет собой стационарный случайный процесс, но уже с отличным от нуля средним значением:

$$\langle J(t)\rangle = \frac{1}{2}\,\langle\xi^2(t)\rangle = \frac{1}{2}\,\psi_\xi(0) = \int_0^{\infty} g_\xi(\nu)\,d\nu. \qquad (5.5.4)$$

Наконец, обычным порядком вводятся *автоковариационная функция и спектральная плотность интенсивности излучения*:

$$\psi_J(\tau) = \langle J(t)J(t+\tau)\rangle - \langle J(t)\rangle^2 = \int_{-\infty}^{\infty} e^{i\cdot 2\pi\nu\tau} g_J(\nu)\,d\nu, \qquad (5.5.5)$$

$$g_J(\nu) = \int_{-\infty}^{\infty} e^{-i\cdot 2\pi\nu\tau} \psi_J(\tau)\,d\tau. \qquad (5.5.6)$$

Для нахождения функции $\psi_J(\tau)$ необходимо вычислить в (5) момент 4-го порядка $\xi(t)$. Эта задача существенно упрощается для случая, когда $\xi(t)$ подчиняется гауссовскому распределению, как это часто бывает ввиду центральной предельной теоремы. При этом

$$\psi_J(\tau) = \frac{1}{2}\,\psi_\xi^2(\tau) \qquad (5.5.7)$$

(Рид, 1962). Согласно (1.3.9) и (1.3.10), преобразование Фурье произведения функций равно свертке их образов. Поэтому из (6), (7) и (2) находим для гауссовского процесса:

$$g_J(\nu) = \frac{1}{2}\int_{-\infty}^{\infty} g_\xi(\nu')\,g_\xi(\nu-\nu')\,d\nu'. \qquad (5.5.8)$$

На практике очень часто рассматривают случай, когда спектральная плотность $g_\xi(\nu)$ не мала лишь в окрестности некоторой частоты ν_0, т.е., как говорят, сигнал $\xi(t)$ является *узкополосным*.

В качестве примера обратимся к вопросу о формировании спектральной линии оптического диапазона ($\nu_0 \sim 5\cdot 10^{14}$Гц) в излучении разреженного газа. Предположим, что в отсутствие столкновений между атомами каждый из них излучал бы неограниченную гармоническую волну $a\cos(2\pi\nu_0 t+\varphi)$, а действие столкновений приводит лишь к случайным сбоям фазы. Тогда поле, создаваемое одним атомом, можно представить в виде

$$\xi_1(t) = a\,\cos[2\pi\nu_0 t+\varphi(t)], \qquad (5.5.9)$$

где a = const и $\varphi(t)$ - стационарный случайный процесс, сохраняющий свое значение в течение промежутков времени случайной длительности θ между столкновениями, а после каждого столкновения принимающий равномерно распределенное на $[0, 2\pi]$ случайное значение. Длина свободного пробега атома экспоненциально распределена, так что плотность

$$p_\theta(\tau) = \frac{1}{\tau_c}\,e^{-\tau/\tau_c}, \qquad 0 \leqslant \tau < \infty, \qquad (5.5.10)$$

где τ_c - среднее время между столкновениями. Таким образом, в данной модели $\xi_1(t)$ представляет собой отрезки синусоиды (цуги) случайной длины.

Процесс характеризуется двумя временными масштабами: периодом $P_0 = \nu_0^{-1}$ волны и временем когерентности τ_c, для разреженного газа значительно превосходящим P_0. Другими словами, имеется *несущая частота* ν_0, в малой окрестности $\Delta\nu \sim \tau_c^{-1}$ которой и сосредоточена спектральная плотность $g_{\xi_1}(\nu)$, содержащая информацию о состоянии газа.

Усредняя $J_1(t) = \xi_1^2(t)/2$ по времени, находим среднюю интенсивность $\langle J_1 \rangle = a^2/4$. Нетрудно получить и точное выражение для $g_{\xi_1}(\nu)$; пренебрегая в нем вкладом максимума при $\nu \simeq -\nu_0$, что справедливо в случае $\tau_c \gg P_0$, имеем для положительных частот:

$$g_{\xi_1}(\nu) = \frac{a^2}{4\pi} \frac{\gamma}{(\nu-\nu_0)^2+\gamma^2}, \qquad (5.5.11)$$

где

$$\gamma = (2\pi\tau_c)^{-1} \ll \nu_0 \qquad (5.5.12)$$

- ширина линии, обусловленная столкновениями. Лоренцовский вид профиля линии (II) сохраняется и при суммировании полей отдельных атомов с естественной заменой a^2 на $A^2 = \sum a_k^2$:

$$g_{\xi}(\nu) = \frac{A^2}{4\pi} \frac{\gamma}{(\nu-\nu_0)^2+\gamma^2}, \qquad \nu \geqslant 0 . \qquad (5.5.13)$$

Средняя интенсивность суммарного поля $\langle J \rangle = A^2/4$. Его значения $\xi(t) = \sum \xi_k(t)$ при большом количестве слагаемых нормально распределены, так что можно найти по формулам (7) и (8) автоковариационную функцию и спектральную плотность интенсивности.

Проведенное обсуждение ясно показывает смысл общего представления узкополосного сигнала в виде

$$\xi(t) = a(t) \cos[2\pi\nu_0 t + \varphi(t)], \qquad (5.5.14)$$

где амплитуда $a(t)$ (ее чаще называют *огибающей*) и фаза $\varphi(t)$ мало изменяются за время $P_0 = \nu_0^{-1}$. Поскольку постоянная времени приемников излучения обычно значительно превосходит

P_0, и потому нас интересуют свойства «медленных» случайных функций $a(t)$ и $\varphi(t)$, представляется естественным усреднить мгновенные характеристики процесса типа $J(t)$ по периоду несущей частоты. Обозначая соответствующее среднее значение через $I(t)$, находим из (3) и (9)

$$I(t) \equiv \frac{1}{P_0}\int_t^{t+P_0} J(t')dt' \simeq \frac{a^2(t)}{4}\,. \qquad (5.5.15)$$

Таким образом, статистические характеристики средней по периоду интенсивности $I(t)$ тесно связаны с аналогичными характеристиками огибающей $a(t)$ узкополосного сигнала. Этот факт и определяет полезность представления (14).

Усредняя первое из равенств (15), получаем

$$\overline{I} \equiv \langle I(t)\rangle = \langle J(t)\rangle = \frac{1}{2}\,\psi_\xi(0), \qquad (5.5.16)$$

т.е. среднее значение сохраняется. Найдем плотность распределения $p_I(x)$ величины $I(t)$ для наиболее важного случая, когда $\xi(t)$ является гауссовским процессом. С этой целью введем наряду с $\xi(t)$ вспомогательный процесс

$$\eta(t) = a(t)\sin[2\pi\nu_0 t + \varphi(t)]. \qquad (5.5.17)$$

Поскольку для узкополосного процесса $a(t)$ и $\varphi(t)$ почти постоянны в течение интервалов времени $\sim P_0$, мы имеем $\eta(t) \simeq \simeq \xi(t-P_0/4)$ и, следовательно, $\eta(t)$ - тоже гауссовский процесс с нулевым средним и той же дисперсией, что и $\xi(t)$. Интенсивность $I(t) = [\xi^2(t)+\eta^2(t)]/4$ пропорциональна сумме квадратов нормальных случайных величин, а мы уже имели повод в связи с (2.3.23) отметить, что такая сумма подчиняется экспоненциальному распределению (Крамер, 1946, п. 18.1).

Итак, плотность распределения усредненной по периоду P_0 интенсивности $I(t)$ равна

$$p_I(x) = (\overline{I})^{-1}e^{-x/\overline{I}},\quad 0 \leqslant x < \infty\,, \qquad (5.5.18)$$

где $\overline{I}$ задана соотношением (16). Этот результат является искомым, он понадобится нам при последующем обсуждении счета фотонов.

Следует подчеркнуть, что экспоненциальное распределение (18) справедливо лишь для некогерентного поля, в то время

как, скажем, для хорошо стабилизированного одномодового лазера интенсивность очень мало флуктуирует около среднего значения $\overline{I}$. Поэтому для излучения лазера можно считать

$$p_I(x) = \delta(x-\overline{I}), \tag{5.5.19}$$

где $\delta(z)$ - функция Дирака (см. Приложение 2).

Распределение фотоотсчетов. Вернемся к обсуждению процесса детектирования фотонов, схематически изображенного на рис. 5.5.1,*а*. Предположим, что скорость фотоионизаций, обусловленных стационарным световым потоком, пропорциональна усредненной по периоду P_0 интенсивности $I(t)$. Именно, пусть при заданном в момент t значении $I(t)$ вероятность того, что число фотоионизаций $\Delta\mathfrak{N}(t)$ в течение малого промежутка времени Δt равно k, есть

$$\Pr[\Delta\mathfrak{N}(t) = k|I(t)] = \begin{cases} 1-\varepsilon I(t)\Delta t, & k = 0, \\ \varepsilon I(t)\Delta t, & k = 1, \\ \text{порядка } (\Delta t)^2, & k > 1. \end{cases} \tag{5.5.20}$$

Постоянная ε, характеризующая чувствительность детектора, пропорциональна квантовому выходу. Мы видим, что в данном случае осуществляются условия (5.1.2), означающие ординарность процесса и отсутствие последействия. Вместе с тем $I(t)$ - стационарный случайный процесс, так что описание фотоионизаций дается дважды стохастическим пуассоновским процессом, рассмотренным в § 5.1. Средняя интенсивность этого процесса $n(t)$ равна теперь средней скорости фотоионизаций $\varepsilon I(t)$, а случайная величина $N(T)$, заданная формулой (5.1.32), превращается в

$$N(T) = \varepsilon \int_0^T I(t)dt \ . \tag{5.5.21}$$

Наконец, вероятность того, что произойдет k фотоионизаций в течение интервала времени длиной T, дается распределением Кокса-Мандела (5.1.35):

$$p_k(T) = \int_0^\infty e^{-x} \frac{x^k}{k!} f_N(x,T)dx, \quad k = 0, 1, \ldots , \tag{5.5.22}$$

где $f_N(x,T)$ - плотность распределения $N(T)$. Рассмотрим некоторые следствия общего результата (22).

Среднее количество фотоотсчетов за время T определяется формулой (5.1.37), т.е.

$$M(T) \equiv \langle \mathfrak{N}(T) \rangle = \varepsilon \overline{I}\, T, \tag{5.5.23}$$

где $\overline{I}$ - средняя интенсивность света (16). Согласно (5.1.40) дисперсия числа фотоотсчетов равна

$$D(T) = M(T) + 2\varepsilon^2 \int_0^T (T-\tau)\phi_I(\tau)\,d\tau\ . \tag{5.5.24}$$

Мы учли здесь то обстоятельство, что автоковариационные функции $n(t)$ и $I(t)$ отличаются множителем ε. В частности, при $\phi_I(\tau) \propto \exp(-|\tau|/\tau_0)$ дисперсия пропорциональна величине (5.1.44).

Последние два соотношения показывают, что для теплового поля излучения дисперсия числа фотоионизаций *больше* среднего их количества, т.е. привычное «пуассоновское» равенство среднего и дисперсии не выполняется. Полезно обсудить несколько конкретных видов поля излучения.

Рассмотрим сначала хорошо стабилизированный лазер. Полагая в соответствии с (19), что его интенсивность фиксирована, для плотности распределения величины (21) получаем $f_N(x,T) = \delta[x-M(T)]$. Подставляя это выражение в (22), находим:

$$p_k(T) = e^{-M(T)}\,\frac{M^k(T)}{k!}\ , \qquad k = 0,\ 1\ldots, \tag{5.5.25}$$

т.е. количество фотоотсчетов подчиняется распределению Пуассона.

Для теплового поля излучения при условии, что время экспозиции T намного меньше времени когерентности τ_c, из (21) следует $N(T) \simeq \varepsilon I(0) T$. С другой стороны, интенсивность $I(0)$ гауссовского поля подчиняется экспоненциальному распределению (18), так что плотность распределения $N(T)$ также имеет экспоненциальную форму:

$$f_N(x,T) \simeq e^{-x/M(T)}\,\frac{1}{M(t)}\ , \qquad 0 \leqslant x < \infty\ . \tag{5.5.26}$$

Из (26) и (22) следует

$$p_k(T) = \frac{[M(T)]^k}{[1+M(T)]^{1+k}}, \qquad k = 0, 1, \ldots \tag{5.5.27}$$

- количество фотоотсчетов подчиняется в этом случае *геометрическому распределению* (см. Приложение 4). Весьма показательно, что геометрическое распределение описывает и совокупность идентичных частиц в одном квантовом состоянии, подчиняющихся статистике Бозе-Зйнштейна. При $T \ll \tau_c$ распределение фотоотсчетов отражает свойства фотонного газа, причем указанное неравенство означает фактически, что подсчитываются фотоны, содержащиеся в одной ячейке фазового пространства (Бертолотти, 1974). Дисперсия числа фотоотсчетов для геометрического распределения дается формулой (П.4.14):

$$D(T) = M(T)[1+M(T)]. \tag{5.5.28}$$

Первое слагаемое справа отвечает классическим частицам, а второе отражает чисто квантовое взаимодействие фотонов вследствие вынужденного излучения (Уэбб, 1972). Мы отмечали в § 2.5, что геометрическое распределение обладает максимальной энтропией среди всех дискретных распределений с заданным средним значением.

В противоположном случае $T \gg \tau_c$ весь интервал T можно разбить на большое количество промежутков длиной τ_c, вклады которых в $N(T)$ независимы. Тогда величина $N(T)$ распределена приблизительно нормально при всяком распределении $I(t)$, причем среднее значение $N(T)$ дается формулой (23), а дисперсия $\sigma^2(T)$ равна второму слагаемому в правой части (24). При $T \to \infty$ отношение $\sigma(T)/M(T) \propto T^{-1/2} \to 0$, т.е. распределение $f_N(x,T)$ стягивается в узкий пик типа δ-функции, а потому, как и для лазерного излучения, распределение фотоотсчетов приближается к пуассоновскому распределению (25).

Таким образом, только в предельном случае, когда время экспозиции T значительно превосходит время когерентности света τ_c, количество фотоотсчетов подчиняется распределению Пуассона.

Инварианты фильтра. Вполне естественно, что вероятностные характеристики фотоотсчетов, заданные распределением

Кокса-Мандела (22), зависят от чувствительности приемника ε. В частности, среднее значение (23), как и ожидалось, пропорционально ε, а дисперсия (24) содержит квадратичное по ε слагаемое. Если бы мы захотели охарактеризовать отклонение от пуассоновской статистики значением индекса дисперсии $\nu \equiv D/M$, то обнаружили бы, что ν линейно изменяется с ε. Однако чаще всего представляют интерес свойства потока излучения, не зависящие от параметров используемого приемника. Так, из (23) и (24) видно, что комбинация первых моментов

$$\frac{D-M}{M^2} \equiv \frac{\nu-1}{M} = \frac{2}{(\bar{I}T)^2}\int_0^T (T-\tau)\psi_I(\tau)\,d\tau \qquad (5.5.29)$$

может рассматриваться в качестве такой характеристики собственно светового потока. Поскольку изменение ε эквивалентно введению некоторого «фильтра», пропускающего каждый из падающих на него фотонов с вероятностью $q < 1$, то можно сказать, что отыскиваются величины, не зависящие от прозрачности фильтра q. Такие величины были названы *инвариантами фильтра* (Пимонов и Теребиж, 1979, 1981).

Заметим, что инвариантность (29) следует из формул (23) и (24), справедливых для произвольного стационарного поля излучения. Поэтому выражение (29) можно использовать для объективного описания всякой стационарной переменности светового потока, например, периодических или случайных вариаций средней интенсивности (см. §§ 5.3 и 5.4). В противоположность этому такие величины как индекс дисперсии $\nu(T)$ следует применять лишь для сравнительного описания статистики фотоотсчетов при разных временах экспозиции T (само по себе значение ν зависит, в частности, от диаметра телескопа и других параметров, входящих в ε). Если речь идет не о статистике фотоотсчетов, а о произвольном потоке событий (скажем, импульсах напряжения в детекторе гравитационных волн), то требование инвариантности не ставится.

В связи с формулой (29) возникает естественный вопрос о существовании других инвариантов фильтра. Соответствующее рассмотрение (Пимонов и Теребиж, 1981) показало, что для потока фотонов с произвольными статистическими характерис-

тиками инварианты можно представить в виде следующей комбинации так называемых *факториальных моментов*:

$$H_m(T) = \frac{\langle \mathfrak{N}(\mathfrak{N}-1)\ldots(\mathfrak{N}-m+1)\rangle}{\langle \mathfrak{N}\rangle^m}, \quad m = 1, 2, \ldots, \qquad (5.5.30)$$

где $\mathfrak{N}(T)$ - случайный процесс фотоионизаций. В частности, $H_1(T) \equiv 1$, а

$$H_2(T) = \frac{\langle \mathfrak{N}(\mathfrak{N}-1)\rangle}{\langle \mathfrak{N}\rangle^2} = \frac{D+M^2-M}{M^2} = \frac{D-M}{M^2} + 1, \qquad (5.5.31)$$

и мы вновь приходим к выводу об инвариантности (29).

Вывод соотношения (30) основывается в работе Пимонова и Теребижа (1981) на воображаемом эксперименте с установкой перед приемником фильтра. Другой вывод нетрудно усмотреть из соотношений (21) и (22). Действительно, величина $N(T)$ лишь множителем ε отличается от $W(T) \equiv \int_0^T I(t)dt$. Плотность распределения последней $f_W(x,T)$ не зависит от ε, а плотность $f_N(x;T) = f_W(x/\varepsilon, T)/\varepsilon$. Подставляя это выражение в (22) и образуя производящую функцию $\{p_k(T)\}$, находим:

$$G(s,T) = \int_0^\infty e^{-\varepsilon z(1-s)} f_W(z,T)dz. \qquad (5.5.32)$$

Отсюда следует, что факториальные моменты, выражающиеся через производные $G(s,T)$ в точке $s = 1$, пропорциональны соответствующим степеням ε, т.е. выражение (30) от ε не зависит.

Свойства квадратичного инварианта $H_2(T)$ и его поведение для различных моделей переменности потока обсуждались нами в §§ 5.3 и 5.4.

§ 5.6. Эффекты мертвого времени. Отношение сигнала к шуму при счете событий

В предыдущих параграфах данной главы предполагалось, что события, поток которых поступает на счетчик, являются «точечными», а счетчик - идеальным, т.е. не вносящим собственных искажений в статистику. Оба эти предположения на практике выполняются с ограниченной точностью. Например, импульсы ФЭУ-79 размываются вследствие дисперсии времени пролета электронов до ширины ~25 нс. Проходя по тракту ре-

гистрации, импульс испытывает дальнейшее уширение, так что можно весь тракт регистрации, включая детектор, охарактеризовать значением *минимального мертвого времени* τ, необходимого для регистрации одного точечного события. Предполагается, что отдельный акт регистрации выводит счетчик из рабочего состояния на время τ, которое считается ниже для простоты фиксированным. Таким образом, предшествующая теория охватывала случаи, когда мертвое время значительно меньше среднего интервала времени между событиями.

Общая теория счетчиков при условии, что входной поток образует пуассоновскую последовательность, разработана Феллером (1948). Ее изложение приводит, в частности, Гольданский и др. (1959) и Баруча-Рид (1960). В первом пункте настоящего параграфа мы ограничимся сводкой основных результатов и необходимыми пояснениями. Типичные для астрономических наблюдений тракты регистрации исследовали Ферни (1977), построивший эмпирическую зависимость числа просчетов от интенсивности поступающего потока, а также Пимонов и Теребиж (1981), сравнивавшие счетчики с мертвым временем продлевающего и непродлевающего типов. Чаще других в литературе рассматривался счет частиц, однако результаты справедливы и для произвольных потоков событий.

Обратимся вновь к рис. 5.5.1, *б*. Ясно, что при $\tau = 0$ средний выходной поток импульсов будет строго пропорционален среднему потоку фотонов, и мы должны получить отсчеты, расположенные вдоль пунктирной прямой. При $\tau > 0$ по мере возрастания потока фотонов все чаще будет случаться наложение импульсов и как следствие - их просчеты. При скорости счета $n \sim \tau^{-1}$ импульсы распределены настолько тесно (образно говоря, временная ось заполнена в среднем равномерно импульсами шириной τ, непосредственно следующими друг за другом), что регистрируемый поток импульсов не возрастает с увеличением потока фотонов.

Таким образом, эффекты мертвого времени тракта регистрации приводят к нелинейности счетчика и ограничивают его динамический диапазон. Казалось бы, изучив влияние просчетов, можно внести соответствующую поправку в результаты измерений и восстановить информацию. Но в нелинейной

области падает отношение сигнала к шуму (Гольданский и др., 1959), так что указанное восстановление отягчено значительной неопределенностью - работа в нелинейной области в принципе неэффективна. Поскольку для многих счетчиков, применяемых в астрономии при исследовании слабых объектов (например, систем типа IPCS Боксенберга, 1975) максимальная допустимая скорость счета ограничена низкими значениями ~1 имп./с на канал, этот факт заслуживает самого пристального внимания.

Уменьшение S/N при значительной доле просчетов тем более примечательно, что флуктуации наблюдаемой скорости счета импульсов в нелинейной области *меньше* пуассоновских при той же средней скорости счета. Не раз случалось, что связанная с этим явлением стабильность отсчетов создавала у наблюдателей уверенность в высокой точности измерений, между тем как действительная точность была низка ввиду существенной неопределенности *оценки* истинного входного потока. Общее определение отношения сигнала к шуму, применимое и к нелинейным счетчикам (Теребиж, 1981а) , рассматривается во втором пункте данного параграфа.

Наконец, в третьем пункте приводятся основные соотношения, описывающие работу счетчика в так называемом режиме N, когда наперед заданным является не время экспозиции, а накопленное количество импульсов (Теребиж, 1983). Такой режим удобен, в частности, при последовательном измерении с примерно одинаковой точностью потоков сильно различающейся интенсивности.

Эффекты мертвого времени. Отличие мертвого времени от нуля приводит к тому, что, во-первых, регистрируются не все поступающие на вход тракта регистрации события и, во-вторых, последовательность импульсов обнаруживает статистическую зависимость даже в том случае, когда на вход поступает пуассоновская последовательность точечных событий. Обозначим среднюю интенсивность последней через n_0, а среднее и дисперсию числа зарегистрированных в интервале $(0, T]$ импульсов шириной τ - соответственно через $M_\tau(T)$ и $D_\tau(T)$.

Задания лишь ширины импульса недостаточно для описания работы счетчика. Обычно приборы, используемые для счета

частиц, с хорошим приближением можно отнести к двум типам: счетчикам с *непродлевающим* и *продлевающим* мертвым временем или соответственно типов I и II. В обоих типах счетчиков поступающие в течение мертвого времени импульсы не регистрируются, однако в счетчиках типа I эти импульсы теряются, не удлиняя мертвое время, тогда как в счетчиках типа II - удлиняют соответственно времени своего прихода (рис. 5.6.1).

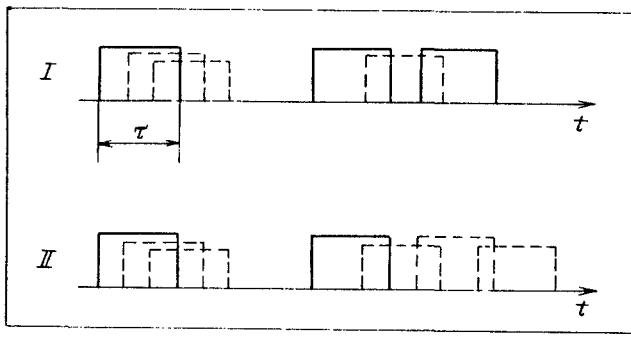


Рис. 5.6.1. Схема регистрации импульсов ширины τ в счетчиках типов I и II. Штриховой линией обозначены пропущенные импульсы

Точная теория постоена Феллером (1948) для произвольного соотношения между временем экспозиции T, шириной импульса τ и средним временем между поступающими событиями n_0^{-1}. Обычно в прикладных задачах можно полагать $T \gg \tau$ и $T \gg n_0^{-1}$ (напомним также, что по предположению T значительно превосходит время когерентности τ_c и входной поток - пуассоновский). При перечисленных условиях для счетчиков типа I среднее и дисперсия числа импульсов определяются выражениями

$$M_\tau(T) = \frac{n_0 T}{1+n_0\tau}, \qquad (5.6.1)$$

$$D_\tau(T) = \frac{n_0 T}{(1+n_0\tau)^3}, \qquad (5.6.2)$$

а для счетчиков типа II -

$$M_\tau(T) = 1 + n_0 T e^{-n_0\tau}, \qquad (5.6.3)$$

$$D_\tau(T) = n_0 T e^{-n_0\tau}\left(1 - 2n_0\tau e^{-n_0\tau}\right). \qquad (5.6.4)$$

Эти соотношения иллюстрируются рис. 5.6.2.

Формулы (1) и (3) можно получить из простых вероятностных соображений. Например, (1) означает, что среднее число пришедших в течение времени T событий $n_0 T$ равно сумме среднего числа зарегистрированных импульсов $M_\tau(T)$ и среднего числа пропущенных импульсов $M_\tau(T)\tau n_0$. В счетчиках типа I при $n_0 \gg \tau^{-1}$ импульсы располагаются «вплотную» друг за другом и число их за время T не может, очевидно, превышать T/τ. Для счетчиков типа II уменьшение реального времени, в течение которого счетчик способен регистрировать поступающие события, начинает преобладать при $n_0 \geqslant \tau^{-1}$, а число зарегистрированных импульсов начинает уменьшаться с увеличением входного потока. В пределе при $n_0 \gg \tau^{-1}$ регистрируется лишь первое из пришедших в $(0, T]$ событий.

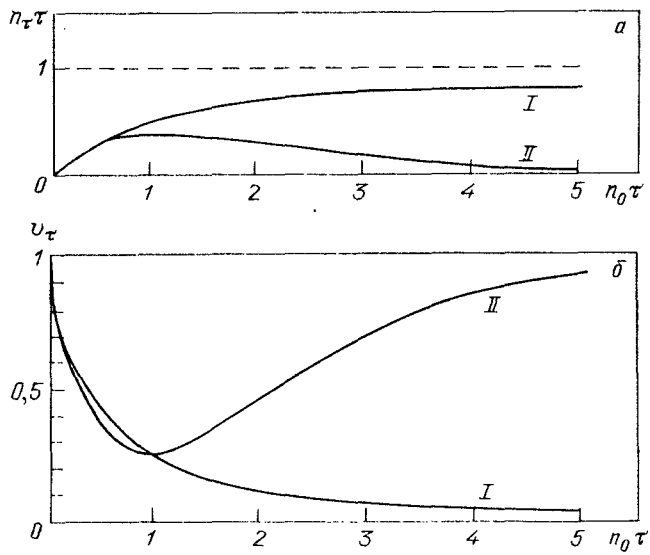


Рис. 5.6.2. Зависимость средней скорости счета (а) и индекса дисперсии (б) от интенсивности поступающего потока событий для счетчиков типов I и II

Поведение дисперсии проще всего характеризовать отношением $V_\tau = D_\tau / M_\tau$, т.е. индексом дисперсии (напомним, что для пуассоновской последовательности $V_0 \equiv 1$). Из (1) - (4) имеем

$$V_\tau = \frac{1}{(1+n_0\tau)^2}, \qquad (5.6.5)$$

$$V_\tau = 1 - 2n_0\tau e^{-n_0\tau} \qquad (5.6.6)$$

соответственно для счетчиков типов I и II (рис. 5.6.2, *б*). Монотонное уменьшение V_τ с ростом n_0 для счетчиков типа I вполне понятно: импульсная последовательность становится все более регулярной, и при $n_0 \gg \tau^{-1}$ всегда регистрируется одно и то же количество импульсов T/τ. Для счетчиков типа II V_τ уменьшается вплоть до $n_0 = \tau^{-1}$, однако в дальнейшем с ростом интенсивности поступающих событий периоды реального мертвого времени становятся столь продолжительными, что последовательные регистрируемые импульсы можно считать независимыми и V_τ вновь приближается к 1.

Выражения (1)-(3) для $M_\tau(T)$ позволяют определить из наблюдений значение минимального мертвого времени τ. Если речь идет о счете фотонов, то интенсивность входного потока n_0 можно изменять, наводя телескоп на *стандартные звезды* (т.е. с известным потоком) различной яркости или, что лучше, диафрагмируя телескоп при наблюдении одной звезды. В области $n_0 \ll \tau^{-1}$ отклонения от линейности, предсказываемые (1) и (3), имеют одинаковый вид: наблюдаемая скорость счета $n_\tau \equiv M_\tau/T \simeq n_0(1-n_0\tau)$. При достаточно большом количестве измерений величин n_0 и n_τ эта зависимость позволяет оценить τ с точностью до процентов. В качестве верхней границы динамического диапазона обычно принимают значение входного потока $n_0^{max} = 0,01\tau^{-1}$, при котором отклонение от линейности достигает 1 %. Например, при $\tau = 50$ нс максимальный доступный поток составляет $\sim 2 \cdot 10^5$ имп./с. Полезно проверить относительное уменьшение дисперсии в области $n_0 \sim \tau^{-1}$. Для грубой оценки можно использовать и тот факт, что наблюдаемая скорость счета n_τ не превосходит τ^{-1} в случае (1) и $(e\tau)^{-1}$ в случае (3).

Отношение сигнала к шуму в нелинейной области. Предполо-

жим, что на вход счетчика поступает пуассоновский поток событий со средней интенсивностью n_0(1/с), а на выходе к потоку импульсов средней интенсивности n, обусловленному входными событиями, добавляется темновой поток n_d. Пусть наблюдения занимают время T; среднее и дисперсия числа входных событий равны $M_0 = D_0 = n_0 T$. Величину M_0 можно рассматривать в качестве «сигнала» S на входе, а $\sqrt{D_0}$ — в качестве «шума» N, так что

$$\left(\frac{S}{N}\right)_{in} = \frac{M_0}{\sqrt{D_0}} = \sqrt{n_0 T} \quad . \tag{5.6.7}$$

При общепринятом подходе сигналом на выходе будет $M(T)$ - среднее накопленное количество «полезных» импульсов, а шумом - стандартное отклонение суммы потоков $\sqrt{D+D_d}$, откуда

$$\left(\frac{S}{N}\right)_{out} = \frac{M(T)}{\sqrt{D+D_d}} \quad . \tag{5.6.8}$$

Применим это определение к описанию *линейного* счетчика, включающего в себя и детектор с квантовым выходом ε (имп.соб). Очевидно, при этом $M(T) = D(T) = nT = \varepsilon n_0 T$, $D_d(T) = M_d(T) = n_d T$, и мы получаем

$$\left(\frac{S}{N}\right)_{out} = \sqrt{\frac{\varepsilon n_0 T}{1+n_d/\varepsilon n_0}} \quad . \tag{5.6.9}$$

Из (7) и (9) видно, что отношение

$$DQE \equiv \left(\frac{S}{N}\right)^2_{out} \Big/ \left(\frac{S}{N}\right)^2_{in} \quad , \tag{5.6.10}$$

называемое *квантовой эффективностью обнаружения* (Detective Quantum Efficiency), определяется главным образом свойствами приемника:

$$DQE = \frac{\varepsilon}{1 + n_d/\varepsilon n_0} \quad . \tag{5.6.11}$$

В отсутствие темнового потока мы получили бы, что DQE просто равна квантовому выходу ε; последнее соотношение показывает, что наличие темнового потока снижает DQE.

Важно подчеркнуть, что определение (10) носит общий ха-

рактер, не связанный обязательно с рассмотренным выше конкретным примером. Параметр DQE описывает эффективность работы детектора и тракта регистрации в *реальных* условиях наблюдения, включающих шум и прочие мешающие факторы.

Итак, принимая (8), мы пришли для линейного приемника к разумным выводам. Нетрудно убедиться, что в нелинейном случае это определение теряет смысл. Предположим для простоты, что темновые импульсы отсутствуют, а приемник и тракт регистрации представляют собой счетчик типа I с квантовым выходом ε и минимальным мертвым временем τ. Тогда (1) и (2) можно записать в виде

$$M_\tau(T) = \frac{\varepsilon n_0 T}{1+\varepsilon n_0 \tau}, \quad D_\tau(T) = \frac{\varepsilon n_0 T}{(1+\varepsilon n_0 \tau)^3}. \qquad (5.6.12)$$

Подстановка этих выражений в (8) дает

$$\left(\frac{S}{N}\right)_{out} = \sqrt{\varepsilon n_0 T(1+\varepsilon n_0 \tau)}, \qquad (5.6.13)$$

так что

$$DQE = \varepsilon(1+\varepsilon n_0 \tau). \qquad (5.6.14)$$

Согласно (13) и (14), отличие мертвого времени от нуля приводит к увеличению как отношения сигнала к шуму, так и квантовой эффективности обнаружения! Разумеется; это не так, и два последних выражения неверны - они отражают лишь то, что последовательность импульсов на выходе счетчика становится с ростом входного потока все более упорядоченной вследствие просчетов. В пределе при $\varepsilon n_0 \gg \tau^{-1}$ отсчеты за время T почти не меняются и равны T/τ, так что шум, формально понимаемый как стандартное отклонение выходного потока импульсов, очень мал, но вместе с тем понятно, что это не говорит о действительно высоком отношении сигнала к шуму.

Итак, мы пришли к необходимости обобщить понятия сигнала и шума. Однозначное обобщение может быть сделано следующим образом (Теребиж, 1981а). Прежде всего обратим внимание на тот факт, что задачей наблюдателя является получение *оценки* $\hat{n}_0$ средней интенсивности входного потока n_0. Будучи функцией наблюдаемых случайных величин, оценка $\hat{n}_0$ сама является

случайной величиной. Хорошая оценка должна быть несмещенной и состоятельной, т.е. ее среднее значение должно быть равно истинному среднему n_0, а дисперсия $D(\hat{n}_0)$ – убывать до нуля с увеличением времени экспозиции. Ясно, что именно среднее значение $\langle\hat{n}_0\rangle$ следует рассматривать в качестве сигнала, а стандартное отклонение оценки $\sqrt{D(\hat{n}_0)}$ – в качестве шума, т.е.

$$\left(\frac{S}{N}\right)_{out} = \frac{\langle\hat{n}_0\rangle}{\sqrt{D(\hat{n}_0)}} . \qquad (5.6.15)$$

Сохраняя определение (10) квантовой эффективности обнаружения, мы имеем также:

$$DQE = \frac{1}{n_0 T}\,\frac{\langle\hat{n}_0\rangle^2}{D(\hat{n}_0)} . \qquad (5.6.16)$$

Применим определения (15) и (16) к рассматривавшемуся выше примеру. Первое из соотношений (12) показывает, что в качестве оценки $\hat{n}_0$ целесообразно принять величину

$$\hat{n}_0 = \frac{\mathfrak{N}}{\varepsilon(T-\tau\mathfrak{N})} , \qquad (5.6.17)$$

где $\mathfrak{N}(T)$ – количество зарегистрированных за время T импульсов. Принимая во внимание, что $\langle\mathfrak{N}\rangle = M_\tau(T)$, $D(\mathfrak{N}) = D_\tau(T)$, и пользуясь правилами приближенного переноса ошибок (Хадсон, 1964, § 9), находим

$$\langle\hat{n}_0\rangle \simeq n_0, \quad D(\hat{n}_0) \simeq \frac{n_0(1+\varepsilon n_0\tau)}{\varepsilon T} . \qquad (5.6.18)$$

Из (15), (16) и (18) следует

$$\left(\frac{S}{N}\right)_{out} = \sqrt{\frac{\varepsilon n_0 T}{1+\varepsilon n_0\tau}} , \quad DQE = \frac{\varepsilon}{1+\varepsilon n_0\tau} . \qquad (5.6.19)$$

Мы видим, что в счетчике с конечным мертвым временем при увеличении интенсивности входного потока отношение сигнала к шуму на выходе и квантовая эффективность обнаружения уменьшаются, как этого интуитивно и следовало ожидать.

Соотношения (15) и (16) имеют общий характер и применимы для описания произвольных приемников, в частности, нелинейных или с неаддитивными шумами.

Отношение сигнала к шуму при накоплении фиксированного количества импульсов. Метод получения оценки блеска какого-либо объекта, обычно используемый в технике счета фотонов, включает два этапа: 1) регистрацию импульсов при одновременном наблюдении объекта и фона (включая темновой ток) в течение некоторого *заданного* промежутка времени T_1; 2) регистрацию фоновых импульсов в течение *заданного* промежутка времени T_2. Обозначим полученные при указанных двух экспозициях количества импульсов соответственно через $\mathfrak{N}_1$ и $\mathfrak{N}_2$. Тогда в качестве оценки средней скорости счета импульсов, обусловленных излучением объекта, принимают величину

$$\hat{n}_* = \frac{\mathfrak{N}_1}{T_1} - \frac{\mathfrak{N}_2}{T_2}. \qquad (5.6.20)$$

Предполагая, что процессы поступления импульсов от объекта и фона являются пуассоновскими процессами со средними интенсивностями соответственно n_* и n_b, нетрудно показать, что $\hat{n}_*$ - несмещенная оценка, т.е. среднее значение $\langle \hat{n}_* \rangle = n_*$, а дисперсия

$$D(\hat{n}_*) = \frac{n_*}{T_1} + \left(\frac{1}{T_1} + \frac{1}{T_2}\right) n_b. \qquad (5.6.21)$$

Применяя формулу (15), находим, что отношение сигнала к шуму в данном случае определяется известным выражением

$$\frac{S}{N} = \frac{\langle \hat{n}_* \rangle}{\sqrt{D(\hat{n}_*)}} = n_* \left[\frac{n_*}{T_1} + \left(\frac{1}{T_1} + \frac{1}{T_2}\right) n_b\right]^{-1/2}. \qquad (5.6.22)$$

Можно доказать, что оценка (20) - наилучшая с точки зрения метода максимального правдоподобия.

Наряду с описанным методом, предполагающим регистрацию случайного количества импульсов в течение фиксированного промежутка времени, в последние годы используется метод, при котором регистрируется некоторое наперед заданное количество импульсов. Величиной, служащей при этом для оценки блеска объекта, является *время накопления заданного числа*

импульсов. Очевидным примером ситуации, когда целесообразно применить второй подход, называемый «режимом N», является сканирование спектра, содержащего яркие относительно континуума эмиссионные линии. Если мы хотим приблизительно сохранить точность измерений в различных участках спектра, то нужно фиксировать накопленное число событий N, и тогда континуум будет измеряться дольше, чем линии.

Итак, задача заключается в нахождении оценки интенсивности объекта и отношения сигнала к шуму при накоплении заданного числа импульсов, т.е. соотношений, аналогичных (20)-(22).

Пусть при наблюдении «объект + фон» счетчик останавливается после регистрации N_1 импульсов, а при наблюдении «фон» - N_2 импульсов. Соответствующие реализации времени экспозиции обозначим через t_1 и t_2 (подчеркнем еще раз, что теперь N_1 и N_2 - фиксированные числа, а t_1 и t_2 - случайные величины). Казалось бы, в качестве оценки интенсивности объекта вновь следует принять выражение типа (20), т.е. $N_1/t_1 - N_2/t_2$. Однако, как показано ниже, эта оценка является смещенной, что неудобно при интерпретации наблюдений слабых объектов.

Для отыскания оценки с нужными свойствами рассмотрим следующую общую задачу. Пусть на счетчик поступает пуассоновский поток импульсов с интенсивностью n. Нужно найти плотность распределения случайной величины θ_N - промежутка времени от начала наблюдения до момента регистрации N-го импульса. Поскольку θ_N есть сумма N независимых случайных величин - промежутков времени между последовательными импульсами, каждая из которых подчиняется экспоненциальному распределению (5.1.14), искомая плотность легко находится применением характеристических функций и оказывается равной

$$f_N(t,n) = \frac{n}{(N-1)!}(nt)^{N-1}e^{-nt}, \quad 0 \leqslant t < \infty . \qquad (5.6.23)$$

Это - известное *гамма-распределение* или *специальное распределение Эрланга* (см., например, Феллер, 1966). Формулы для моментов имеют следующий вид:

$$\left\langle \frac{1}{\theta_N} \right\rangle = \frac{n}{N-1}, \quad \left\langle \frac{1}{\theta_N^2} \right\rangle = \frac{n^2}{(N-1)(N-2)}, \quad N > 2 , \qquad (5.6.24)$$

Обратимся теперь к рассмотренной выше процедуре наблюдений. Согласно методу максимального правдоподобия, в качестве оценок параметров n_* и n_b следует принять такие величины $\hat{n}_*$ и $\hat{n}_b$, при которых вероятность получить наблюдавшиеся значения t_1 и t_2 максимальна. Указанная вероятность пропорциональна

$$L = f_{N_1}(t_1, n_*+n_b)\ f_{N_2}(t_2, n_b)\ . \qquad (5.6.25)$$

Подставляя (23) в (25) и приравнивая логарифмические производные L нулю, находим

$$\hat{n}_* = \frac{N_1}{t_1} - \frac{N_2}{t_2}\ , \quad \hat{n}_b = \frac{N_2}{t_2}\ . \qquad (5.6.26)$$

При помощи (24) определяются средние значения:

$$\langle\hat{n}_*\rangle = \frac{N_1}{N_1-1}(n_*+n_b) - \frac{N_2}{N_2-1}\ n_b,$$
$$\langle\hat{n}_b\rangle = \frac{N_2}{N_2-1}\ n_b\ . \qquad (5.6.27)$$

Таким образом, обе оценки (27) являются смещенными, причем особенно нежелательным является то обстоятельство, что смещение $\hat{n}_*$ зависит от интенсивности фона. В качестве несмещенных оценок следует принять величины

$$\hat{n}_* = \frac{N_1-1}{t_1} - \frac{N_2-1}{t_2}\ , \quad \hat{n}_b = \frac{N_2-1}{t_2}\ . \qquad (5.6.28)$$

Вновь обращаясь к (24), находим следующее выражение для дисперсии $\hat{n}_*$:

$$D(\hat{n}_*) = \frac{(n_*+n_b)^2}{N_1-2} + \frac{n_b^2}{N_2-2}\ , \qquad (5.6.29)$$

где предполагается $N_1 > 2$, $N_2 > 2$. Наконец, отношение сигнала к шуму для «режима N» равно

$$\frac{S}{N} = \left[\frac{1}{N_1-2}\left(1+\frac{n_b}{n_*}\right)^2 + \frac{1}{N_2-2}\left(\frac{n_b}{n_*}\right)^2\right]^{-1/2}\ . \qquad (5.6.30)$$

Помимо заданных наперед чисел N_1 и N_2 оно определяется

лишь отношением $\delta \equiv n_b/n_*$. Для часто встречающегося случая, когда $N_1 = N_2 = N$, имеем из (30)

$$\frac{S}{N} = \sqrt{\frac{N-2}{1+2\delta(1+\delta)}}\,, \qquad (5.6.31)$$

где предполагается $N > 2$.

Соотношения (28) - (31) являются искомыми. Они дают полное описание наблюдений при накоплении заданного количества импульсов. В предельном случае, когда N_1 и N_2 велики, можно положить $N_1 - 2 \simeq (n_* + n_b)t_1$, $N_2 - 2 \simeq n_b t_2$. При этом выражение (30) совпадает с (22), т.е. «режим N» и «режим T» асимптотически эквивалентны.

Выражение (31) ясно показывает, что для сравнительно ярких источников (когда $n_* \gg n_b$) отношение сигнала к шуму равно интуитивно ожидаемому значению $\sim\sqrt{N}$. При наблюдениях же слабых объектов, создающих поток порядка фонового, действительная точность измерений задается полными выражениями (30) и (31).

Очевидно, надлежащее изменение контекста позволяет использовать найденные выше соотношения не только при оценивании яркости источников света, о котором для определенности говорилось в данном пункте, но и при решении других задач теории счета событий.

ЗАКЛЮЧЕНИЕ

Изучение временных рядов имеет достаточно длительную историю, и тем не менее эта область продолжает интенсивно развиваться. Так, столетие с момента появления работы Лагранжа (1873) ознаменовалось разработкой весьма эффективного подхода, основанного на информационных критериях. В наши дни значительные усилия направляются на развитие теории импульсных процессов. Внимание, которое уделяется анализу временных последовательностей, обусловлено, главным образом, требованиями практики: к таким рядам фактически относится всякая совокупность измерений, где имеет значение порядок отсчетов. Современные компьютеры позволяют реализовать методы, ранее представлявшие чисто теоретический интерес. Наконец, отчасти разнообразие появившихся в 60 - 80-е годы методов связано с недостаточной информированностью исследователей.

Вообще говоря, разнообразие методов представляется вполне естественным, поскольку эффективность их неодинакова в различных условиях. Вместе с тем отсутствие стандартной методики создает известные проблемы при оценивании надежности выводов. Большое число работ, связанных с обнаружением лож-

ных зависимостей, периодичностей и т.п., заставляет еще раз подчеркнуть общее положение: мы вправе не рассматривать результат, если при его выводе не соблюдены определенные *необходимые* условия. К таковым прежде всего следует отнести *полноту* описания и проверку *устойчивости* результатов. Первое условие наиболее ясно выражено М.Кендаллом (1976): «... исследователь временных рядов не может быть пуристом. Что он, безусловно, может сделать,- это привести первоначальные данные, над которыми он работал, и дать недвусмысленные объяснения относительно того, как он их обрабатывал, для того чтобы каждый, кто не согласен с этим, смог провести собственный анализ». Что касается второго необходимого условия, то оно имеет в виду варьирование методики анализа, последовательный выбор различных частей изучаемого ряда, моделирование близких численных примеров и пр.; доверия заслуживают лишь те выводы, которые *сохраняются* при разумном изменении подхода.

Конечно, все перечисленные выше условия не гарантируют сами по себе корректности анализа. Исследователь должен «чувствовать» надежность экспериментальных данных, выбрать подходящий способ их рассмотрения и, не в последнюю очередь, опираться на здравый смысл.

Приложение 1. Значения некоторых сумм

$$\sum_{k=0}^{N-1}\cos(2kx) = \frac{\cos[(N-1)x]\sin(Nx)}{\sin x}, \qquad (П.1.1)$$

$$\sum_{k=1}^{N-1}\sin(2kx) = \frac{\sin[(N-1)x]\sin(Nx)}{\sin x}, \qquad (П.1.2)$$

$$\sum_{k=0}^{N-1}p^k\cos(2kx)=\frac{1-p\cos(2x)-p^N\cos(2Nx)+p^{N+1}\cos[2(N-1)x]}{1-2p\cos(2x)+p^2}, \qquad (П.1.3)$$

$$\sum_{k=1}^{N-1}p^k\sin(2kx)=\frac{p\sin(2x)-p^N\sin(2Nx)+p^{N+1}\sin[2(N-1)x]}{1-2p\cos(2x)+p^2}, \qquad (П.1.4)$$

$$\sum_{k=0}^{N-1}\cos^2(kx) = \frac{N+1}{2} + \frac{\sin[(N-1)x]\cos(Nx)}{2\sin x}, \qquad (П.1.5)$$

$$\sum_{k=1}^{N-1}\sin^2(kx) = \frac{N-1}{2} - \frac{\sin[(N-1)x]\cos(Nx)}{2\sin x}, \qquad (П.1.6)$$

$$\sum_{k=-n}^{n}\cos(2kx) = \frac{\sin[(2n+1)x]}{\sin x}, \qquad (П.1.7)$$

$$\sum_{k=-n}^{n}\left(1-\frac{|k|}{n}\right)^2 = \frac{2}{3}n + \frac{1}{3n}. \qquad (П.1.8)$$

Приложение 2. Основные формулы с δ-функциями Дирака

Функция Дирака $\delta(x)$ относится к классу обобщенных функций. На уровне физической строгости ее можно представить таким образом, что $\delta(x) = 0$ для всех $x \neq 0$, а при $x = 0$ эта функция обладает столь сильной сингулярностью, что сходится интеграл

$$\int_{-\infty}^{\infty}\delta(x)dx = 1. \qquad (П.2.1)$$

Сами по себе формулы с δ-функциями имеют символический смысл; они становятся строгими при условии, что $\delta(x)$ стоит под знаком интеграла.

Имея в виду сделанные оговорки, рассмотрим $\delta(x)$ как предел некоторой вспомогательной функции $\delta(x,Z)$ при $Z \to \infty$. В качестве $\delta(x,Z)$ можно принять, например, функцию

$$\delta(x,Z) = \int_{-Z}^{Z} e^{i\cdot 2\pi xz} dz = \frac{\sin(2\pi xZ)}{\pi x}, \qquad (\text{П.2.2})$$

так что

$$\delta(x) = \lim_{Z\to\infty} \frac{\sin(2\pi xZ)}{\pi x}, \quad \delta(-x) = \delta(x), \qquad (\text{П.2.3})$$

и справедливо интегральное представление

$$\delta(x) = \int_{-\infty}^{\infty} e^{i\cdot 2\pi xz} dz. \qquad (\text{П.2.4})$$

Из (П.2.3) нетрудно показать, что для всякой регулярной при $x = 0$ функции $f(x)$ имеет место формула замены аргумента:

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)dx = f(a). \qquad (\text{П.2.5})$$

В теории временных рядов часто используется представление $\delta(x)$ в виде суммы:

$$\delta(x) = \nu_0 \sum_{n=-\infty}^{\infty} e^{i\cdot 2\pi n\nu_0 x}. \qquad (\text{П.2.6})$$

Суперпозиция функций

$$\delta[f(x)] = \sum_{s} |f'(x_s)|^{-1} \delta(x-x_s), \qquad (\text{П.2.7})$$

где суммирование распространяется на корни x_s уравнения $f(x) = 0$, расположенные в рассматриваемом интервале. В частности,

$$\delta(ax) = \frac{1}{|a|}\delta(x). \qquad (\text{П.2.8})$$

Другая, помимо (П.2.2), конкретная форма вспомогательной

функции $\delta(x,Z)$ использовалась в § 1.4:

$$\delta(x,Z) = \frac{1}{Z}\left[\frac{\sin(\pi xZ)}{\pi x}\right]^2 \xrightarrow[Z\to\infty]{} \delta(x). \qquad (П.2.9)$$

Как функция (П.2.2), так и (П.2.9) удовлетворяют условию нормировки

$$\int_{-\infty}^{\infty} \delta(x,Z)\,dx = 1, \qquad (П.2.10)$$

следующему из формул (П.3.5).

Приложение 3. Свойства функций sinc(x), $\phi_N(x)$ и tr(x)

Вспомогательные элементарные функции $\mathrm{sinc}(x)$ и $\phi_N(x)$ определяются соотношениями

$$\mathrm{sinc}(x) = \frac{\sin(\pi x)}{\pi x}, \quad -\infty < x < \infty, \qquad (П.3.1)$$

$$\phi_N(x) = \frac{\sin(\pi Nx)}{N\sin(\pi x)}, \quad -\infty < x < \infty, \quad N = 1, 2, \ldots \qquad (П.3.2)$$

Для первой из этих функций имеем (см., в частности, Брейсуэлл, 1965)

$$\mathrm{sinc}(0) = 1, \quad \mathrm{sinc}(-x) = \mathrm{sinc}(x), \qquad (П.3.3)$$

$$\int_{-\infty}^{\infty} \mathrm{sinc}(x)\,dx = \int_{-\infty}^{\infty} \mathrm{sinc}^2(x)\,dx = 1, \qquad (П.3.4)$$

$$\int_{-\infty}^{\infty} e^{-i\cdot 2\pi\nu x}\,\mathrm{sinc}(x)\,dx = r_1(\nu) = \begin{cases} 1, & |\nu| \leqslant 1/2, \\ 0, & |\nu| > 1/2, \end{cases} \qquad (П.3.5)$$

$$\int_{-\infty}^{\infty} e^{-i\cdot 2\pi\nu x}\,\mathrm{sinc}^2(x)\,dx = \begin{cases} 1-|\nu|, & |\nu| \leqslant 1, \\ 0, & |\nu| > 1. \end{cases} \qquad (П.3.6)$$

График $\mathrm{sinc}(x)$ представлен на рис. 1.3.1.

Функция $\phi_N(x)$ удовлетворяет равенствам:

$$\phi_N(0) = 1, \quad \phi_N(-x) = \phi_N(x), \qquad (П.3.7)$$

$$\phi_N(x \pm 2k) = \phi_N(x), \quad k = 0, 1, 2, \ldots, \qquad (П.3.8)$$

$$\phi_N(x \pm k) = \begin{cases} -\phi_N(x), & \text{если } N \text{ - четное, } k \text{ - нечетное,} \\ \phi_N(x) & \text{в других случаях.} \end{cases} \qquad (П.3.9)$$

Последнее соотношение показывает, что $|\phi_N(x)|$ - периодическая функция с периодом, равным 1, т.е.

$$|\phi_N(x\pm k)| = |\phi_N(x)|, \quad k = 0, 1, 2, \ldots \qquad (\text{П.3.10})$$

Интеграл от $\phi_N(x)$ по периоду равен

$$I \equiv \int_{-1/2}^{1/2} \phi_N(x)dx = \begin{cases} \frac{1}{N}, & \text{при } N \text{ нечетном,} \\ \frac{4}{\pi N}\left[1 - \frac{1}{3} + \frac{1}{5} - \ldots + \frac{(-1)^{N/2-1}}{N-1}\right], & \text{при } N \text{ четном.} \end{cases} \qquad (\text{П.3.11})$$

В случае $N \gg 1$ мы имеем $I \simeq N^{-1}$ и при четных N. Функцию $\phi_N^2(x)$ часто называют ядром Фейера. Его нормировка имеет вид

$$\int_{-1/2}^{1/2} \phi_N^2(x)dx = \frac{1}{N}. \qquad (\text{П.3.12})$$

Ядро Фейера при различных значениях N представлено на рис. 2.1.1.

При обсуждении влияния тренда на периодограмму удобно ввести функцию

$$\mathrm{tr}(x) \equiv \frac{1}{\pi}\frac{d}{dx}\,\mathrm{sinc}(x) = \frac{1}{\pi x}\left(\cos\pi x - \frac{\sin\pi x}{\pi x}\right). \qquad (\text{П.3.13})$$

При малых значениях x имеем

$$\mathrm{tr}(x) \simeq -\frac{\pi x}{3}\left(1 - \frac{\pi^2 x^2}{10} + \ldots\right). \qquad (\text{П.3.14})$$

Основные максимумы $\mathrm{tr}^2(x)$ расположены при $x_{max} \simeq \pm 0,6626$ и $\mathrm{tr}^2(x_{max}) \simeq 0,19025$. График $\mathrm{tr}^2(x)$ представлен на рис. 3.5.1.

Приложение 4. Характеристики случайных величин

В данном приложении мы перечислим использующиеся в книге свойства некоторых вероятностных распределений.

Всякая случайная величина ξ задается 1) совокупностью значений, которые она может принимать, и 2) вероятностями осуществления этих значений. В классическом примере с бросанием игральной кости случайную величину можно определить, скажем, как «число очков на верхней грани кости». Тогда возможными значениями ξ будут целые числа 1, 2, ..., 6. Для

симметричной кости соответствующие вероятности равны $p_1 = p_2 = \ldots = p_6 = 1/6$, но для реальной кости распределение вероятностей отлично от равномерного.

Типичные вероятностные задачи связаны с нахождением распределений некоторых сложных случайных величин, если вероятности всех элементарных событий заданы. Например, при независимом бросании трех игральных костей можно интересоваться распределением случайной величины η, определенной как «сумма очков на верхних гранях», причем распределения очков для каждой кости ξ_1, ξ_2 и ξ_3 должны быть заданы наперед.

Обозначим возможные значения дискретной случайной величины ξ посредством $x_1, x_2, \ldots, x_n$ (где n может быть и бесконечно большим), а соответствующие им вероятности - посредством $p_1, p_2, \ldots, p_n$. Иногда случайную величину и ее возможные значения обозначают одинаковым образом, однако во избежание путаницы к этому прибегают лишь в исключительных случаях.

Вероятностное распределение всегда нормировано к 1:

$$\sum_{k=1}^{n} p_k = 1. \qquad (\text{П.4.1})$$

Среди численных характеристик случайных величин важнейшими являются среднее значение и дисперсия:

$$\langle \xi \rangle = \sum_{k=1}^{n} x_k p_k, \quad D_\xi \equiv \sigma_\xi^2 = \sum_{k=1}^{n} (x_k - \langle \xi \rangle)^2 p_k. \qquad (\text{П.4.2})$$

Наиболее часто встречаются *целочисленные* случайные величины, принимающие значения $k = 0, 1, 2, \ldots$ с некоторыми вероятностями $p_0, p_1, p_2, \ldots$ При обращении с такими величинами весьма эффективным средством является *производящая функция вероятностей* (ПФВ), определенная следующим образом:

$$G(s) = \sum_{k=0}^{\infty} p_k s^k, \quad -1 \leqslant s \leqslant 1, \quad G(1) = 1. \qquad (\text{П.4.3})$$

При сложении независимых случайных величин их производящие функции перемножаются. Знание ПФВ позволяет определить все

характеристики случайной величины, в частности,

$$\langle\xi\rangle = G'(1),\quad D_\xi = G''(1) + G'(1)[1 - G'(1)]. \qquad (\text{П.}4.4)$$

Для непрерывной случайной величины ξ задается вероятность $p(x)dx$ ее попадания в бесконечно малый интервал $[x, x+dx]$, где $-\infty < x < \infty$. Функция $p(x)$ называется *плотностью распределения вероятностей*. Аналогами (П.4.1) и (П.4.2) являются формулы

$$\int_{-\infty}^{\infty} p(x)dx = 1,\quad \langle\xi^n\rangle = \int_{-\infty}^{\infty} x^n p(x)dx,\quad n = 1, 2, \ldots, \qquad (\text{П.}4.5)$$

$$D_\xi \equiv \sigma_\xi^2 = \langle\xi^2\rangle - \langle\xi\rangle^2.$$

Вместо $G(s)$ вводится характеристическая функция

$$\varphi(t) = \langle e^{it\xi}\rangle = \int_{-\infty}^{\infty} e^{itx} p(x)dx; \qquad (\text{П.}4.6)$$

при сложении независимых случайных величин их характеристические функции перемножаются. Знание $\varphi(t)$ позволяет найти $p(x)$ путем обращения преобразования Фурье; первые моменты

$$\langle\xi\rangle = -i\varphi'(0),\quad D_\xi = -\varphi''(0) + [\varphi'(0)]^2. \qquad (\text{П.}4.7)$$

Биномиальное распределение. Отличны от нуля вероятности

$$p_k = C_n^k p^k (1-p)^{n-k},\quad k = 0, 1, \ldots, n;\ 0 \leqslant p \leqslant 1, \qquad (\text{П.}4.8)$$

где $C_n^k = n/k!(n-k)!$ - биномиальные коэффициенты. ПФВ определяется выражением

$$G(s) = [1 - (1-s)p]^n. \qquad (\text{П.}4.9)$$

Первые моменты равны

$$\langle\xi\rangle = np,\quad D_\xi = np(1-p). \qquad (\text{П.}4.10)$$

В частности, биномиальным является распределение случайной величины ξ, определенной как число успехов в серии из n независимых испытаний, если в отдельном испытании вероятность успеха равна p.

Распределение Пуассона можно рассматривать как предел биномиального распределения при $p \to 0$, $n \to \infty$, $np = a = \text{const}$, т.е. как распределение числа успехов при очень

большом количестве попыток реализовать маловероятное событие. Из (П.4.8) - (П.4.10) нетрудно найти для пуассоновского распределения:

$$p_k = e^{-a}\frac{a^k}{k!}, \quad k = 0, 1, \ldots; \; a > 0,$$
$$G(s) = e^{-(1-s)a}, \quad \langle\xi\rangle = D_\xi = a. \tag{П.4.11}$$

Геометрическое распределение задано вероятностями

$$p_k = \frac{a^k}{(1+a)^{1+k}}, \quad k = 0, 1, \ldots; \quad a > 0. \tag{П.4.12}$$

Производящая функция

$$G(s) = \frac{1}{1+(1-s)a}, \tag{П.4.13}$$

первые моменты

$$\langle\xi\rangle = a, \quad D_\xi = a(1+a). \tag{П.4.14}$$

Если обозначить $p \equiv (1+a)^{-1}$, то формулу (П.4.12) можно переписать в виде $p_k = (1-p)^k p$, позволяющем интерпретировать ξ как длину серии до первого успеха в ряде независимых испытаний, в каждом из которых вероятность успеха равна p.

Прямоугольное распределение имеет случайная величина, принимающая с равной вероятностью любое значение из конечного промежутка. Обычно рассматривают равномерно распределенную на промежутке [0, 1] случайную величину α, плотность которой равна

$$p(x) = \begin{cases} 1, & 0 \leqslant x \leqslant 1, \\ 0, & x < 0, \quad x > 1; \end{cases} \tag{П.4.15}$$

характеристическая функция -

$$\varphi(t) = (e^{it} - 1)/it. \tag{П.4.16}$$

Нетрудно найти, что

$$\langle\alpha\rangle = \frac{1}{2}, \quad D_\alpha = \frac{1}{12}. \tag{П.4.17}$$

Случайная величина ξ с плотностью распределения $p_\xi(x)$,

$a \leqslant x \leqslant b$ связана с α соотношением

$$\int_a^{\xi} p_{\xi}(x)dx = \alpha, \qquad (\text{П.4.18})$$

позволяющим численно моделировать сложный закон распределения ξ по реализациям равномерной случайной величины α.

Экспоненциальное распределение имеет плотность

$$p(x) = \begin{cases} a^{-1}e^{-x/a}, & 0 \leqslant x < \infty, \\ 0, & x < 0, \end{cases} \qquad (\text{П.4.19})$$

характеристическую функцию

$$\varphi(t) = (1-iat)^{-1} \qquad (\text{П.4.20})$$

и первые моменты

$$\langle\xi\rangle = a, \quad D_{\xi} = a^2. \qquad (\text{П.4.21})$$

В статистической литературе, посвященной временным рядам, это распределение часто выступает как χ^2-распределение с двумя степенями свободы. При численном моделировании экспоненциальной случайной величины ξ можно пользоваться соотношением

$$\xi = a \ln\alpha, \qquad (\text{П.4.22})$$

где α - случайная величина, равномерно распределенная на [0, 1]. Последняя формула следует из (П.4.18), если учесть, что α и $1-\alpha$ одинаково распределены.

Нормальное распределение задается двухпараметрической плотностью

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad -\infty < x < \infty. \qquad (\text{П.4.23})$$

Характеристическая функция

$$\varphi(t) = e^{iat-\frac{\sigma^2 t^2}{2}}, \qquad (\text{П.4.24})$$

первые моменты

$$\langle\xi\rangle = a, \quad D_{\xi} = \sigma^2. \qquad (\text{П.4.25})$$

Если α_1 и α_2 - две независимые равномерно распределенные на

[0, 1] случайные величины, то

$$\xi_1 = \sqrt{-2\ln\alpha_1}\,\sin(2\pi\alpha_2),$$
$$\xi_2 = \sqrt{-2\ln\alpha_1}\,\cos(2\pi\alpha_2) \qquad (\text{П. 4. 26})$$

- две независимые нормальные величины со средним 0 и дисперсией 1.

χ^2-*распределение*. Пусть каждая из независимых случайных величин ξ_k нормальна с нулевым средним и единичной дисперсией. Тогда случайная величина

$$\chi_n^2 = \xi_1^2 + \xi_2^2 + \ldots + \xi_n^2 \qquad (\text{П. 4. 27})$$

имеет плотность распределения

$$p_n(x) = 2^{-\frac{n}{2}}\Gamma^{-1}\left(\frac{n}{2}\right)x^{\frac{n}{2}-1}\,e^{-\frac{x}{2}},\quad 0 \leqslant x < \infty, \qquad (\text{П. 4. 28})$$

и $p_n(x) = 0$ при $x < 0$. Параметр n называют числом степеней свободы χ^2-распределения. Характеристическая функция χ_n^2 равна

$$\varphi_n(t) = (1 - 2it)^{-n/2}, \qquad (\text{П. 4. 29})$$

первые моменты

$$\langle\chi_n^2\rangle = n, \quad D(\chi_n^2) = 2n. \qquad (\text{П. 4. 30})$$

Таблицы χ^2-распределения имеются во многих руководствах по статистике.

Приложение 5. Процентные точки для отношения максимального значения периодограммы к среднему

Таблица П.5.1

N	$X_{N,q}$		N	$X_{N,q}$	
	q=0,05	q=0,01		q=0,05	q=0,01
6	1,950	1,990	142	6,967	8,428
8	2,613	2,827	162	7,122	8,601
10	3,072	3,457	182	7,258	8,750
12	3,419	3,943	202	7,378	8,882
14	3,697	4,331	302	7,832	9,372
16	3,928	4,651	402	8,147	9,707
18	4,125	4,921	502	8,389	9,960
20	4,297	5,154	602	8,584	10,164
22	4,450	5,358	702	8,748	10,334
32	5,019	6,103	802	8,889	10,480
42	5,408	6,594	1002	9,123	10,721
52	5,701	6,955	1202	9,313	10,916
62	5,935	7,237	1402	9,473	11,079
82	6,295	7,663	1602	9,612	11,220
102	6,567	7,977	1802	9,733	11,344
122	6,785	8,225	2002	9,842	11,454

Приложение 6. Процентные точки для критерия Аббе

Таблица П.6.1

N	$A_{N,q}$			N	$A_{N,q}$		
	q=0,05	0,01	0,001		q=0,05	0,01	0,001
4	0,3902	0,3128	0,2949	32	0,7177	0,6089	0,4963
5	0,4102	0,2690	0,2080	33	0,7216	0,6141	0,5027
6	0,4451	0,2808	0,1817	34	0,7256	0,6193	0,5090
7	0,4680	0,3070	0,1848	35	0,7292	0,6242	0,5150
8	0,4912	0,3314	0,2018	36	0,7328	0,6290	0,5208
9	0,5121	0,3544	0,2210	37	0,7363	0,6337	0,5265
10	0,5311	0,3759	0,2408	38	0,7396	0,6381	0,5319
11	0,5482	0,3957	0,2598	39	0,7429	0,6425	0,5373
12	0,5638	0,4140	0,2778	40	0,7461	0,6467	0,5425
13	0,5778	0,4309	0,2949	41	0,7491	0,6508	0,5475
14	0,5908	0,4466	0,3112	42	0,7521	0,6548	0,5524
15	0,6027	0,4611	0,3266	43	0,7550	0,6587	0,5571
16	0,6137	0,4746	0,3413	44	0,7576	0,6622	0,5616
17	0,6237	0,4872	0,3552	45	0,7603	0,6659	0,5660
18	0,6330	0,4989	0,3684	46	0,7628	0,6693	0,5701
19	0,6417	0,5100	0,3809	47	0,7653	0,6727	0,5743
20	0,6498	0,5203	0,3926	48	0,7676	0,6757	0,5781
21	0,6574	0,5301	0,4037	49	0,7698	0,6787	0,5817
22	0,6645	0,5393	0,4142	50	0,7718	0,6814	0,5853
23	0,6713	0,5479	0,4241	51	0,7739	0,6842	0,5887
24	0,6776	0,5562	0,4334	52	0,7759	0,6869	0,5922
25	0,6836	0,5639	0,4423	53	0,7779	0,6896	0,5955
26	0,6893	0,5713	0,4509	54	0,7799	0,6924	0,5989
27	0,6946	0,5784	0,4591	55	0,7817	0,6949	0,6020
28	0,6996	0,5850	0,4670	56	0,7836	0,6974	0,6051
29	0,7046	0,5915	0,4748	57	0,7853	0,6999	0,6083
30	0,7091	0,5975	0,4822	58	0,7872	0,7024	0,6114
31	0,7136	0,6034	0,4895	59	0,7891	0,7049	0,6145
				60	0,7906	0,7071	0,6174

Приложение 7. Числовые данные

Таблица П.7.1. Ряд А ($N = 50$)

f_k	ξ_k	y_k	f_k	ξ_k	y_k
1,000	-0,383	0,617	0,650	+0,330	0,980
0,222	-0,314	-0,092	0,029	-0,293	-0,264
0,003	-0,056	-0,053	0,636	-1,874	-1,238
0,888	+0,279	1,167	0,366	+0,287	0,653
0,085	-0,334	-0,249	0,007	+0,179	0,186
0,076	-0,135	-0,059	0,985	+0,125	1,110
0,620	+0,267	0,887	0,163	-0,184	-0,021
0,025	+0,091	0,116	0,007	-0,181	-0,174
0,723	-0,633	0,090	0,804	-0,119	0,685
0,342	-0,789	-0,447	0,057	-0,172	-0,115
0,006	-0,210	-0,204	0,214	+0,263	0,477
0,976	-0,793	0,183	0,517	+0,127	0,644
0,148	+0,243	0,391	0,016	+0,082	0,098
0,010	-0,528	-0,518	0,959	+0,986	1,945
0,778	+0,080	0,858	0,262	-0,681	-0,419
0,051	+0,441	0,492	0,004	+0,129	0,133
0,273	-0,339	-0,066	0,926	+0,109	1,035
0,488	+0,256	0,744	0,104	-0,142	-0,038
0,014	-0,732	-0,718	0,038	-0,105	-0,067
0,989	+0,749	1,738	0,679	+0,270	0,949
0,242	+0,581	0,823	0,033	+0,069	0,102
0,003	+0,461	0,464	0,548	+0,564	1,112
0,908	-0,534	0,374	0,392	-0,251	0,141
0,094	-0,239	-0,145	0,008	+0,103	0,111
0,054	+0,661	0,715	0,992	-0,069	0,923

Таблица П.7.2. Числа Вольфа (N = 210)

Год	R	Год	R	Год	R	Год	R
1749	80,9	1802	45,0	1855	6,7	1908	48,5
1750	83,4	1803	43,1	1856	4,3	1909	43,9
1751	47,7	1804	47,5	1857	22,8	1910	18,6
1752	47,8	1805	42,2	1858	54,8	1911	5,7
1753	30,7	1806	28,1	1859	93,8	1912	3,6
1754	12,2	1807	10,1	1860	95,7	1913	1,4
1755	9,6	1808	8,1	1861	77,2	1914	9,6
1756	10,2	1809	2,5	1862	59,1	1915	47,4
1757	32,4	1810	0,0	1863	44,0	1916	57,1
1758	47,6	1811	1,4	1864	47,0	1917	103,9
1759	54,0	1812	5,0	1865	30,5	1918	80,6
1760	62,9	1813	12,2	1866	16,3	1919	63,6
1761	85,9	1814	13,9	1867	7,3	1920	37,6
1762	61,2	1815	35,4	1868	37,3	1921	26,1
1763	45,1	1816	45,8	1869	73,9	1922	14,2
1764	36,4	1817	41,1	1870	139,1	1923	5,8
1765	20,9	1818	30,4	1871	111,2	1924	16,7
1766	11,4	1819	23,9	1872	101,7	1925	44,3
1767	37,8	1820	15,7	1873	66,3	1926	63,9
1768	69,8	1821	6,6	1874	44,7	1927	69,0
1769	106,1	1822	4,0	1875	17,1	1928	77,8
1770	100,8	1823	1,8	1876	11,3	1929	65,0
1771	81,6	1824	8,5	1877	12,3	1930	35,7
1772	66,5	1825	16,6	1878	3,4	1931	21,2
1773	34,8	1826	36,3	1879	6,0	1932	11,1
1774	30,6	1827	49,7	1880	32,3	1933	5,7
1775	7,0	1828	62,5	1881	54,3	1934	8,7
1776	19,8	1829	67,0	1882	59,7	1935	36,1
1777	92,5	1830	71,0	1883	63,7	1936	79,7
1778	154,4	1831	47,8	1884	63,5	1937	114,4
1779	125,9	1832	27,5	1885	52,2	1938	109,6
1780	84,8	1833	8,5	1886	25,4	1939	88,8
1781	68,1	1834	13,2	1887	13,1	1940	67,8
1782	38,5	1835	56,9	1888	6,8	1941	47,5
1783	22,8	1836	121,5	1889	6,3	1942	30,6
1784	10,2	1837	138,3	1890	7,1	1943	16,3
1785	24,1	1838	103,2	1891	35,6	1944	11,0
1786	82,9	1839	85,8	1892	73,0	1945	33,2
1787	132,0	1840	63,2	1893	84,9	1946	92,6
1788	130,9	1841	36,8	1894	78,0	1947	151,5
1789	118,1	1842	24,2	1895	64,0	1948	136,2
1790	89,9	1843	10,7	1896	41,8	1949	134,7
1791	66,6	1844	15,0	1897	26,2	1950	83,9
1792	60,0	1845	40,1	1898	26,7	1951	69,3
1793	46,9	1846	61,5	1899	12,1	1952	31,5
1794	41,0	1847	98,5	1900	9,5	1953	13,9
1795	21,3	1848	124,3	1901	2,7	1954	4,4
1796	16,0	1849	95,9	1902	5,0	1955	38,0
1797	6,4	1850	66,5	1903	24,4	1956	141,7
1798	4,1	1851	64,5	1904	42,0	1957	189,9
1799	6,8	1852	54,2	1905	63,5	1958	184,8
1800	14,5	1853	39,0	1906	53,8		
1801	34,0	1854	20,6	1907	62,0		

Таблица П.7.3. Кривая блеска квазара 3С 273 (N = 292). В столбцах последовательно даны средние за 100-дневные интервалы времени оценки звездной величины, начиная с JD = 2410480 (Кункель, 1967)

12,83	12,88	12,80	12,55	12,62	12,49	12,83	13,10
12,80	12,91	12,79	12,53	12,67	12,58	12,87	13,12
12,77	12,93	12,78	12,59	12,65	12,66	12,88	13,11
12,75	12,97	12,77	12,66	12,63	12,76	12,88	13,04
12,72	12,99	12,76	12,72	12,61	12,76	12,89	13,01
12,69	13,02	12,75	12,76	12,64	12,52	12,84	12,99
12,67	13,05	12,74	12,74	12,79	12,56	12,73	12,97
12,64	13,04	12,72	12,72	12,98	12,59	12,64	12,96
12,61	13,01	12,70	12,70	13,05	12,58	12,76	12,94
12,59	12,99	12,69	12,68	13,10	12,54	12,83	12,93
12,56	12,91	12,67	12,68	13,07	12,49	12,83	12,91
12,53	12,81	12,71	12,69	12,98	12,52	12,83	12,89
12,51	12,71	12,74	12,69	12,93	12,61	12,82	12,87
12,48	12,63	12,75	12,70	12,95	12,58	12,79	12,80
12,45	12,62	12,82	12,70	12,98	12,57	12,76	12,71
12,42	12,71	12,86	12,71	13,03	12,68	12,80	12,71
12,39	12,80	12,91	12,72	13,05	12,79	12,88	12,73
12,37	12,89	12,87	12,73	13,05	12,77	12,88	12,96
12,38	12,92	12,83	12,73	13,11	12,74	12,88	12,84
12,41	12,95	12,77	12,74	13,12	12,74	12,87	12,91
12,44	12,99	12,75	12,73	12,95	12,76	12,85	12,98
12,47	13,02	12,78	12,71	12,92	12,77	12,83	13,00
12,51	13,05	12,81	12,67	12,90	12,77	12,80	13,00
12,54	13,08	12,85	12,65	12,91	12,77	12,84	13,00
12,57	13,11	12,84	12,65	12,91	12,80	12,94	13,00
12,60	13,03	12,78	12,81	12,89	12,82	12,93	12,97
12,59	12,93	12,71	12,98	12,87	12,83	12,92	12,89
12,58	12,84	12,65	13,05	12,78	12,83	12,94	12,96
12,57	12,89	12,67	12,89	12,73	12,88	12,96	13,03
12,57	12,88	12,69	12,80	12,72	12,96	12,83	13,02
12,63	12,82	12,72	12,71	12,74	13,00	12,75	13,02
12,70	12,75	12,75	12,62	12,74	13,01	12,82	13,01
12,76	12,74	12,76	12,60	12,70	13,07	12,89	12,93
12,81	12,75	12,72	12,67	12,67	13,09	12,99	
12,82	12,76	12,68	12,71	12,60	13,03	13,06	
12,83	12,77	12,64	12,64	12,56	12,95	13,07	
12,85	12,79	12,59	12,62	12,49	12,87	13,08	

Таблица П.7.4. Кривая блеска двухмодовой цефеиды BK Cen (N = 49) согласно Леотта-Жанин (1967). Моменты t_k = JD Hel -2438000

t_k	ΔV_k	t_k	ΔV_k	t_k	ΔV_k
813,48	0,699	843,27	0,283	876,43	0,614
814,49	0,346	849,37	0,113	877,38	0,825
814,59	0,309	855,38	0,529	877,43	0,731
815,48	0,338	856,27	0,265	883,32	0,721
815,55	0,358	857,38	0,724	883,38	0,700
818,48	0,392	861,36	0,783	886,40	0,760
818,57	0,404	861,48	0,785	904,21	0,317
821,48	0,386	862,26	-0,153	907,22	0,274
823,54	0,689	870,39	0,783	908,20	0,731
823,61	0,686	870,45	0,720	920,19	0,326
824,29	-0,017	871,25	0,356	930,21	0,681
829,27	0,665	873,35	0,569	931,20	0,674
830,41	0,349	873,38	0,654	946,12	0,683
832,28	0,700	874,26	0,599	947,21	0,547
832,52	0,788	875,28	0,129	948,29	0,325
840,27	0,071	876,38	0,608	950,24	0,741
841,58	0,782				

Таблица П.7.5. Оценки блеска звезды CF Cas согласно Бердникову (1986). Моменты t_k = JD Hel -2445000

t_k	m_k	t_k	m_k
509,4296	10,911	686,1601	11,001
512,4335	11,331	687,1445	11,193
513,4296	11,396	688,1913	11,336
514,4256	10,860	689,2381	11,229
649,3554	11,338	690,1601	10,830
656,3984	10,910	691,1875	11,059
660,3046	11,053	692,1367	11,237
665,3476	10,938	693,1445	11,367
666,2617	10,919	694,1679	11,216
660,2772	11,295	695,1756	10,855
676,2578	10,981	701,1756	11,100
677,2109	11,173	705,1131	10,895
679,2187	11,350	706,1210	11,104
683,2147	11,331	707,1250	11,275

СПИСОК ЛИТЕРАТУРЫ

Абраменко В.И., Рачковский Д.Н., 1983// Изв. КрАО. -Т.66.- С.71.

Агекян Т.А., 1974; Теория вероятностей для астрономов и физиков.- М.: Наука.

Ахманов С.А., Дьяков Ю.Е., Чиркин А.С., 1981; Введение в статистическую радиофизику и оптику.- М.: Наука.

Больцман Л., 1956; Лекции по теории газов.- М.: Гостехиздат.

Белоконь Е.Т., 1989; в печати.

Бердников Л.Н., 1986// Переменные звезды.-Т.22.-С.369.

Бернштейн И.Н., Шварцман В.Ф., 1983; не опубликовано.

Бирюков В.В., Теребиж В.Ю., 1989.

Большев Л.Н., Смирнов Н.В., 1965; Таблицы математической статистики.-М.: Наука.

Бор Г., 1934; Почти периодические функции.- М.: ГТТИ.

Брандт З.,1975; Статистические методы анализа наблюдений.- М.: Мир.

Бэкон Д.У., Брокховен Л.Х., 1986// Статистические методы для ЭВМ.- М.: Наука, гл. 15.

Вайнштейн Л.А., Зубаков В.Д., 1960; Выделение сигналов на фоне случайных помех.-М.: Сов. Радио.

Василенко Г.И., Тараторин А.М.,1986; Восстановление изображений.-М.: Радио и связь.

Ватутин В.А., Телевинова Т.М., Чистяков В.П., 1985; Вероятностные методы в физических исследованиях.- М.: Наука.

Вязаницын В.П., Гневышев М.Н., Добровольский О.В., Крат В.А., Марков А.В., Молчанов А.П., Соболев В.М., Шаронов

В.В., 1964; Курс астрофизики и звездной астрономии.- Т.3.-М.: Наука.

Гаген-Торн В.А., 1985// Астрофизика.-Т.22.-С.449.

Гаген-Торн В.А., 1987// Астрофизика.-Т.26.-С.415.

Гельфонд А.О., 1967; Исчисление конечных разностей.- М.: Наука.

Гнеденко Б.В., 1988; Курс теории вероятностей.- М.: Наука.

Гольданский В.И., Куценко А.В., Подгорецкий М.И., 1959; Статистика отсчетов при регистрации ядерных частиц.- М.: Физматгиз.

Гончарский А.В., Черепащук А.М., Ягола А.Г., 1985; Некорректные задачи астрофизики.- М.: Наука.

Демченко В.И., Курчаков А.В., 1983; О реставрации изображений в астрономии.- Алма-Ата: Астрофиз. инст.

Ермаков С.М., Михайлов Г.А., 1982; Статистическое моделирование.- М.: Наука.

Журбенко И.Г., 1982; Спектральный анализ временных рядов.- М.: Изд. МГУ.

Журбенко И.Г., 1987; Анализ стационарных и однородных случайных систем.- М.: Изд. МГУ.

Зельдович Я.Б., 1964// Докл. АН СССР.-Т.155.-С.67.

Иваненко Д.Д., Соколов А.А., 1949; Классическая теория поля.- М.: Гостехиздат.

Колмогоров А.Н., 1941// Изв. АН СССР, сер. мат.- №5.-С.3.

Колмогоров А.Н., 1941// Вестник МГУ, сер. мат.- №2.- С. 40.

Конева Е.С., 1988// Автоматика и телемеханика.- №6.- С.3.

Кули Дж.У., Льюис П.А.У., Уэлч П.Д., 1986// Статистические методы для ЭВМ.- М.: Наука.

Курочкин Н.Е., 1964// Астрон. циркуляр.- №279.

Левин Б.Р., 1968; Теоретические основы статистической радиотехники.- М.: Сов. Радио.

Линник Ю.В., 1958; Метод наименьших квадратов и основы теории обработки наблюдений.- М.: Физматгиз.

Литвиненко О.Н., 1974; Основы радиооптики.- Киев: Техника.

Логинов В.П., 1973// Зарубеж. радиоэлектроника.- №4.- С.73.

Лютый В.М., 1989; в печати.

Майер А.Г., Леонтович Е.А., 1934// Докл. АН СССР.-Т.4.-С.353.

Малахов А.Н., 1968; Флуктуации в автоколебательных систе-

мах.- М.: Наука.

Малышев В.И., 1979; Введение в экспериментальную спектроскопию.- М.: Наука.

Мандельштам Л.И., 1972; Лекции по теории колебаний.- М.: Наука.

Никифоров И.В., 1983; Последовательное обнаружение изменения свойств временных рядов.- М.: Наука.

Озерной Л.М., *Чертопруд В.Е.*, 1966// Астрон. журн.-Т.43.-С.20.

Пельт Я.Э., 1975; Препринт Тартуской астрофиз. обс., №5.

Пельт Я.Э., 1980; Частотный анализ астрономических временных рядов.- Таллинн; Валгус.

Пимонов А.А., *Теребиж В.Ю.*, 1979// Сообщ.Спец. астрофиз. обс.-Т.25.-С.17.

Пимонов А.А., *Теребиж В.Ю.*, 1981// Астрофиз. иссл. (Изв. Спец. астрофиз. обс.).-Т.13.-С.62.

Плахотниченко В.Л., 1983// Сообщ. Спец. астрофиз. обс.-Т.38.- С.29.

Пустыльник Е.И., 1968; Статистические методы анализа и обработки наблюдений.- М.: Наука.

Рачковский Д.Н., 1985// Изв. Крымской астрофизич. обс.-Т.71.-С.25.

Рытов С.М., 1953; Теория электрических флуктуаций и теплового излучения.- М.: Изд. АН СССР.

Рытов С.М., 1976; Введение в статистическую радиофизику. Часть 1. Случайные процессы.- М.: Наука.

Рытов С.М., *Левин М.Л.*, 1967; Теория равновесных тепловых флуктуаций в электродинамике.- М.: Наука.

Серебренников М.Г., *Первозванский А.А.*, 1965; Выявление скрытых периодичностей.- М.: Наука.

Солодов А.В., *Солодов А.А.*, 1988; Статистическая динамика систем с точечными процессами.-М.: Наука.

Сороко Л.М., 1981// Физ. элем. частиц и ат. ядра.- Т.12.-С. 754.

Теребиж В.Ю., 1976// Астрон. циркуляр.-№906.-С.3.

Теребиж В.Ю., 1981а// Астрон. журн.-Т. 58.-С.221.

Теребиж В.Ю., 1981б// Астрон. журн.-Т. 58.-С.437.

Теребиж В.Ю., 1983// Тр. ГАИШ.-Т. 53.-С. 188.

Теребиж В.Ю., 1990// Астрофизика.-Т.32.-С.327.

Теребиж В.Ю., Теребиж А.В., Бирюков В.В., 1989// Астрофизика.-Т.31.-С.75.

Тихонов А.Н., Арсенин В.Я., 1979; Методы решения некорректных задач.- М.: Наука.

Фихтенгольц Г.М., 1960; Курс дифференциального и интегрального исчисления, т. III.- М- Л.: Физматгиз.

Харвиг В., Кальв П., 1975// Астрон. циркуляр.- №876.-С.3.

Харкевич А.А., 1962; Спектры и анализ.- М.: Физматгиз.

Шварцман В.Ф., 1977// Сообщ.Спец.астрофиз. обс.-Т. 19.-С.5.

Шварцман В.Ф., Бескин Г.М., Гершберг Р.Е., Плахотниченко В.Л., Пустильник Л.А., 1988// Письма в астрон. журн.-Т. 14.-С.233.

Шеннон К., 1963; Работы по теории информации и кибернетике.- М.: ИЛ.

Яглом А.М., Яглом И.М., 1973; Вероятность и информация.- М.: Наука.

Ables J.G., 1974// Astron. Astrophys. Suppl.-V.15.-P.383.

Airy G.B., 1835// Trans. Cambr. Phil. Soc.- V.5.- P.283.

Akaike H., 1969a// Ann. Inst. Stat. Math.- V.21.- P.243.

Akaike H., 1969b// Ann. Inst. Stat. Math.- V.21.- P.407.

Akaike H., 1970// Ann. Inst. Stat. Math.- V.22.- P.203.

Allen C.W., 1973// Astrophysical Quantities.- Univ. of London, Athlone Pr.; перевод: Аллен К.У., 1977; Астрофизические величины.- М.: Мир.

Andersen N., 1974// Geophysics.- V.39.-P.69.

Anderson R.L., 1942// Ann. Math. Statist.- V.13.- P.1.

Anderson R.L., Anderson T.W., 1950// Ann. Math. Statist.- V.21.- P.59.

Anderson T.W., 1971; The Statistical Analysis of Time Series.- New York: Wiley; перевод: Андерсон Т., 1976; Статистический анализ временных рядов.- М.: Мир.

Andreasen G.K., Petersen J.O., 1987// Astron. Astrophys.- V.180.-P.129.

Andronov I.L., 1988, Astron. Nachr.-V.309.-P.121.

Angione R.J., Smith H.J., 1985//Astron. J.-V.90.-P.2474.

Bailey S.I., 1901// Harv.Ann-V.38.-P.228.

Barbieri L., Romano G., di Serego A.S., Zambon M., 1977//

Astron. Astrophys.-V.59.-P.419.

Barning F.J.M., 1963// Bull. Astr. Inst. Neth.-V.17.-P.22.

Bartlett M.S., 1946; J.Roy. Stat. Soc., B 8.-P.27.

Bartlett M.S., 1950// Biometrica.-V.37.-P.1.

Bartlett M.S., 1955; An Introduction to Stochastic Processes.- Cambridge Univ. Press; перевод: Бартлетт М.С., 1958; Введение в теорию случайных процессов.-М.:ИЛ.

Berry P.A.M., 1987// Vistas in Astronomy.-V.30.-P.2,97.

Berryman J.G., 1978// Geophysics.-V.43.-P.1384.

Bertolotti M., 1974// Photon Correlation and Light Beating Spectroscopy/ Eds.Cummins H.Z. and Pike E.R.- New York; London: Plenum Press.

Bharucha-Reid A.T., 1960; Elements of the Theory of Markov Processes and Their Applications. -New York: Mc Graw--Hill; перевод: Баруча-Рид А.Т., 1969; Элементы теории марковских процессов и их приложения.-М.: Наука.

Blackman R.B., Tukey J.W., 1959; The Measurement of Power Spectra.- New York: Dower.

Bloomfield P., 1976; Fourier Analysis of Time Series: an Introduction.- New York: Wiley.

Boksenberg A., 1975// Image Processing Techniques in Astronomy/ Ed. by C. de Jager and H. Nieuwenhuizen.- Boston: Reidel.

Boldt E.A., Holt S.S., Serlemitsos P.J., 1971// Astrophys. J.-V.164, L9.

Bopp B.W., Evans D.S., Laing J.D., Deeming T.J., 1970// Monthly Not. RAS.-V.147.-P.355.

Born M., Wolf E., 1964; Principles of Optics.- London: Pergamon Press; перевод: Борн М., Вольф Э., 1970; Основы оптики.- М.: Наука.

Bourret R.C., 1960// Nuovo Cim.-V.18.-P.347.

Box G.E.P., Jenkins G.M., 1970; Time Series Analysis. Forecasting and Control.-San Francisco: Holden-Day; перевод: Бокс Дж., Дженкинс Г., 1974; Анализ временных рядов. Прогноз и управление. Вып.1.-М.: Мир.

Bracewell R.N., 1965; The Fourier Transform and its Applications.-New York: Mc Graw-Hill.

Bracewell R.N., 1983// J. Opt. Soc. Am.-V.73.-P.1832.

Bracewell R.N., 1984// Proc. IEEE.-V.72.-P.1010.

Bracewell R.N., Roberts J.A., 1954// Austr.J. Phys.-V.7.-P.615.

Bradley J.V., 1968; Distribution-Free Statistical Tests.- New York: Prentice-Hall.

Brault J.W., White O.R., 1971// Astron. Astrophys.-V.13.-P.169.

Brillinger D.R., 1975; Time Series. Data Analysis and Theory.- Holt, Rinehart and Winston; перевод: Бриллинджер Д., 1980; Временные ряды. Обработка данных и теория.- М.: Мир.

Brown J.L., 1986// Proc. IEEE.-V.74.-P.1596.

Brown R.H., Twiss R.Q., 1958// Proc. Roy. Soc.-A 248, No.1253., P.199, 222.

Buccheri R., Ozel M.E., Sacco B., 1987// Astron. Astrophys. -V.175.-P.353.

Burg J.P., 1967// Paper presented at the 37-th Ann. Int. Meeting.- Oklahoma City.

Burg J.P., 1968// Paper presented at Advanced Study Inst. on Signal Processing, NATO, Enschede, Netherlands.

Burg J.P., 1972// Geophysics.-V.37.-P.375.

Burke E.W., Rolland W.W., Boy W.R., 1970// J.R. Astr. Soc. Canada.-V.64.-P.353.

Burki G., Maeder A., Rufener F., 1978// Astron. Astrophys.-V.65.-P.363.

Burki G., Rufener F., 1978// Astron. Astrophys. -V.70.-P.105.

Campbell N., 1909a// Proc. Cambr. Phil. Soc.- V.15.-P.117.

Campbell N., 1909b// Proc. Cambr. Phil. Soc.- V.15.-P.310.

Chandrasekhar S., 1943; Stochastic Problems in Physics and Astronomy// Rev. Mod. Phys.- V.15.-P.1. перевод: Чандрасекар С., 1947; Стохастические проблемы в физике и астрономии.- М.: ИЛ.

Chatfield C., 1975; The Analysis of Time Series: Theory and Practice.- New York: Chapman and Hall.

Colgate S.A., 1967// Astrophys. J.- V.150.-P.163.

Colgate S.A., Colvin J.D., Petschek A.G., 1975// Astrophys. J. Lett.- V.197, L105.

Collura A., Maggio A., Sciortino S., Serio S., Vaiana G.S., Rosner R., 1987// Astrophys. J.- V.315, P.1.-P.340.

Cooley J.W., Tukey J. W., 1965// Math. Comp.- V.19.-P.297.

Cox D.R., 1955// J. Roy. Stat. Soc.- V.17, B.-P.129.

Cox D.R., Lewis P.A.W., 1966; The Statistical Analysis of Series of Events.- New York: Wiley; перевод: Кокс Д., Льюис П., 1969; Статистический анализ последовательностей событий.- М.: Мир.

Craddock J.M., 1967// The Statistician.- V.17.-P.257.

Cramer H., 1946; Mathematical Methods of Statistics.- Princeton: Univ. Press.; перевод: Крамер Г., 1975; Математические методы статистики.- М.: Мир.

Crow E.L., Davis F.A., Maxfield M.W., 1960; Statistics Manual.- New York: Dower.

Cummins H.Z., Pike E.R. (eds), 1974; Photon Correlation and Light Beating Spectroscopy.- New York; London: Plenum Press.

Currie R.G., 1973// Astrophys. Sp. Sci.- V.20.-P.509.

Davenport W.B., Root W.L., 1958; An Introduction to the Theory of Random Signals and Noise.- New York; Toronto; London: Mc Graw-Hill; перевод: Давенпорт В.Б., Рут В.Л., 1960; Введение в теорию случайных сигналов и шумов.- М.: ИЛ.

Davis H.T., 1941; The Analysis of Economic Time Series.- Cowles Comm. Res. Econ., Bloomington, Indiana, Monog. No. 6.

Deeming T.J., 1970// Astron. J.- V.75.-P.1027.

Deeming T.J., 1975a// Astrophys. Sp. Sci.- V.36.-P.137.

Deeming T.J., 1975b// Astrophys. Sp. Sci.- V.42.-P.257.

Deeter J.E., Boynton P.E., 1982// Astrophys. J.- V.261.-P.337.

Deeter J.E., 1984// Astrophys. J.- V.281.-P.482.

Delache P., Scherrer P.H., 1983// Nature.- V.306.-P.651.

Diaconis P., Efron B., 1983// Scientific American.- V.248, No. 5.

Dreher J.W., Roberts D.H., Lehar J., 1986// Nature.-V.320.-P.239.

Dumont Th., Morguleff N., Rutily B., Terzan A., 1978//

Astron. Astrophys.-V.69. -P.65.

Du Puy D.L., Hoffman G.A., 1985// Comm. IAPPP.- No.20.-P.1.

Du Puy D.L., 1986// Comm. IAPPP.- No.24.- P.26.

Dutta P., Horn P.M., 1981// Rev. Mcd. Phys.- V.53.-P.497.

Duvall T.L., Jr., Harvey J.W., 1984// Nature.- V.310.-P.19.

Dworetsky M.M., 1983// Monthly Notices Roy. Astron. Soc.- V.203.-P.917.

Eadie W.T., Drijard D., James F.E., Roos M., Sadoulet B., 1971; Statistical Methods in Experimental Physics.- North-Holland: Dordrecht; перевод: Статистические методы в экспериментальной физике, 1976.- М.: Атомиздат.

Ebert K., Ederer H., 1985; Computeran-Wendungen in der Chemie.- Weinheim; перевод: Эберт К., Эдерер Х., 1988; Компьютеры. Применение в химии.- М.: Мир.

Ehrenfest P., Ehrenfest T., 1912// Enzyclopädie der Math. Wiss.- Bd. 4.

Elst E.W., 1972// Astron. Astrophys.- V.17.-P.148.

Enslein K., Ralston A., Wilf H.S. (Eds)., 1977; Statistical Methods for Digital Computers.- New York: Wiley; перевод: Энслейн К., Рэлстон Э., Уилф Г.С. (ред)., 1986; Статистические методы для ЭВМ.- М.: Наука.

Fahlman G.G., Ulrych T.J., 1975// Astrophys. J. -V.201.- P.277.

Fahlman G.G., Ulrych T.J., 1976// Astrophys. J.- V.209.- P.663.

Fahlman G.G., Ulrych T.J., 1982// Monthly Notices Roy. Astron. Soc.- V.199.-P.53.

Fath E.A., 1908// Lick Obs. Bull.- V.5.-P.71.

Faulkner D.J., 1977// Astrophys. J.- V.216.-P.49.

Feller W., 1966; An Introduction to Probability Theory and Its Applications, v.I,II.- New York; London; Sydney: Wiley; перевод: Феллер В., 1967; Введение в теорию вероятностей и ее приложения, т.I,II.-М.: Мир.

Feller W., 1948; Courant Anniversary Volume, Interscience Publishers, New York, p.105.

Fellgett P., 1972// Astrophys. Sp. Sci.- V.16.-P.437.

Fernie J.D., 1977// Publ. Astron. Soc. Pacific.- V.88.-P.969.

Ferraz-Mello S., 1981// Astron. J.- V.86.-P.619.

Feynman R.P., Leighton R.B., Sands M., 1963; The Feynman Lectures on Physics.- Pato Alto, London: Addison-Wesley. перевод: Фейнман Р., Лейтон Р., Сэндс М., 1967; Феймановские лекции по физике.- М.: Мир.

Feynman R.P., 1985// QED - The Strange Theory of Light and Matter.- Princeton: Univ. Press; перевод: Фейнман Р., 1988; КЭД. Странная теория света и вещества.- М.: Наука.

Field G.B., 1964// Astrophys. J.- V.140.-P.1434.

Fisher R.A., 1924// Proc. Int. Math. Congr. Toronto.- P.805.

Fisher R.A., 1929// Proc. Roy. Soc., Ser. A.- V.125.-P.54.

Fisher R.A., 1937; The Design of Experiments.- Edinburgh, London.

Fisher R.A., 1941; Statistical Methods for Research Workers.-Edinburgh, London.

Fitch W.S., 1967// Astrophys. J.- V.148.-P.481.

Fitch W.S. (ed.), 1976; Multiple Periodic Variable Stars// Proc. Int. Astr. Union Coll.- No.29.- Budapest.

Fletcher R., Powell M.J.D., 1974// Math. Comp.- V.28.- P.1067.

Forman W., Jones C., Tananbaum H., 1976// Astrophys. J.- V.208.-P.849.

Fourgere P.F., 1975// EOS Trans. Amer. Geophys. Union.- V.56.-P.1054.

Fourgere P.F., Zawalick E.J., Radoski H.R., 1976// Phys. of the Earth and Planetary Interiors.- V.12.-P.201.

Frasca S., 1985// Data Anal. Astron. Proc. Int. Workshop, Erice, May 28 - June 4, 1984.- New York, London.-P.75.

Frieden B.R., 1979// Picture Processing and Digital Filtering d. T.S. Huang - Springer.-P.179-249.

Frieden B.R., 1985// Proc. IEEE.- V.73.-P.1764.

Fuller W., 1976// Introduction to Statistical Time Series.- New York: Wiley.

Gentleman W.M., 1973// J. Inst. Math. Appl.- V.12.-P.329.

Gentleman W.M., 1974// Appl. Stat.- V.23.-P.448.

Gibson A.I., Harrison A.B., Kirkman I.W., Lotts A.P., Macrae J.H., Orford K.J., Turver K.E., Walmsley W.,

1982// Nature.- V.296.-P.833.

Goldman S., 1953; Information Theory.- London: Constable and Co; перевод: Голдман С., 1957; Теория информации.- М.: Мир.

Gray D.F., 1976// The Observations and Analysis of Stellar Photospheres.- New York: Wiley; перевод: Грей Д., 1980; Наблюдения и анализ звездных фотосфер.- М.: Мир.

Gray D.F., Desikachary K., 1973// Astrophys. J.- V.181.-P.523.

Groth E.J., 1975a// Astrophys. J. Suppl.- V.29.-P.285.

Groth E.J., 1975b// Astrophys. J.Suppl.- V.29.- P.431, 443, 453.

Gull S.F., Skilling J., 1983// IAU/URSI Symp. on Imaging.- Sydney.

Hagen J.G., 1921// Die Veranderlichen Sterne.-Bd. I.-S.622.

Hamming R.W., 1962; Numerical Methods for Scientists and Engeneers.- New York: Mc Graw-Hill; перевод: Хемминг Р.В., 1968; Численные методы для научных работников и инженеров.- М.: Наука.

Hamming R.W., 1977; Digital Filters.- New York: Prentice-Hall.; перевод: Хемминг Р.В., 1980; Цифровые фильтры.- М.: Сов. Радио.

Hannan E.J., 1960; Time Series Analysis.- New York: Wiley; перевод: Хеннан Э., 1964; Анализ временных рядов.- М.: Наука.

Hannan E.J., 1970; Multiple Time Series.- New York: Wiley; перевод: Хеннан Э., 1974; Многомерные временные ряды.- М.: Мир.

Hart B.I., 1942a// Ann Math. Stat.- V.13.-P.207.

Hart B.I., 1942b// Ann. Math. Stat.- V.13.-P.445.

Heck A., Manfroid J., Mersch G., 1985// Astron. Astrophys. Suppl.- V.59.-P.63.

Henning H.M., Scherrer P.H., 1986// CSSA-ASTRO-86-43, Nov.

Hills J.G., 1975// Nature.- V.254.-P.295.

Hills J.G., 1978// Monthly Notices Roy. Astron. Soc.- V.182.-P.517.

Hoffmeister C., 1914// Astron. Nachr.- V.196.-P.399.

Högbom J.A., 1974// Astron. Astrophys. Suppl.-V.15.-P.417.

Hooge F.N., Kleinpenning T.G.M., Vandamme L.K.J., 1981// Rep. Progr. Phys.- V.44.-P.479.

Horman M., 1980; Masters Thesis.- Univ. Liege.

Horne J.H., Baliunas S.L., 1986// Astrophys. J.- V.302.-P.757.

Hudson D.J., 1964; Statistics. Lectures on Elementary Statistics and Probability.- Geneva; перевод: Хадсон, 1970; Статистика для физиков.- М.: Мир.

Jaynes E.T., 1982// Proc. IEEE.- V.70.-P.939.

Jenkins G.M., Watts D.G., 1968; Spectral Analysis and Its Applications.-San Francisco: Holden Day; перевод: Дженкинс Г., Ваттс Д., Вып.1, 1971; Вып.2, 1972; Сектральный анализ и его приложения.- М.: Мир.

Jeudy L.M.A., 1982// Astrophys. Sp. Science.-V.81.-P.275.

Jurkevich I., 1971// Astrophys. Sp. Science.-V.13.-P.154.

Jurkevich I., 1972// Astrophys. J. Lett.-V.172, L29.

Kanasewich E.R., 1981; Time Sequence Analysis in Geophysics.-Univ. of Alberta Press; перевод: Канасевич Э.Р., 1985; Анализ временных последовательностей в геофизике.- М.: Недра.

Kano Y., Wolf E., 1962// Proc. Phys. Soc.- V.80.-P.1273.

Kendall M., 1976; Time-Series.- London: Griffin and comp.; перевод: Кендалл М., 1981; Временные ряды.-М.: Финансы и статистика.

Kendall M.G., Stuart A., 1966; The Advanced Theory of Statistics, Vol.3, Design and Analysis, and Time-Series.- London: Griffin and comp.; перевод: Кендалл М., Стьюарт А., 1976; Многомерный статистический анализ и временные ряды.- М.: Наука.

Kirk B.L., Rust B.W., van Winkle W., 1979; Time Series Analysis by the Maximum Entropy Method.- Environmental Sciences Division Publ. No. 1220, Oak Ridge National Lab.

Kovacs G., 1980// Astrophys. Sp. Science.- V.69.-P.485.

Kovacs G., 1981// Astrophys. Sp. Science.- V.78.-P.175.

Kuhn J.R., 1982// Astron. J.- V.87.-P.196.

Kunkel W.E., 1967// Astron. J.- V.72.-P.1341.

Kurtz D.W., 1983// Comiss. 27 IAU Inf. Bull Var. Stars.-

No.2285.

Kurtz D.W., 1985// Monthly Notices Roy. Astron. Soc.- V.213.-P.773.

Lagrange J.L., 1873// Oeuvres VI.-P.507.

Lacoss R.T., 1971// Geophysics.-V.36.-P.661.

Lafler J., Kinman T.D., 1965// Astrophys. J. Suppl.-V.11.-P.216.

Lampton M., Margon B., Bowyer S., 1976// Astrophys. J.-V.208.-P.177.

Lanczos C., 1956; Applied Analysis.- New York: Prentice-Hall; перевод: Ланцош К., 1961; Практические методы прикладного анализа.- М.: Физматгиз.

Lawrence A., Watson M.G., Pounds K.A., Elvis M., 1987// Nature.- V.325.-P.694.

Leahy D.A., 1987// Astron. Astrophys.- V.180.-P.275.

Leahy D., Darbro W., Elsner R., Weisskopf M., Sutherland P., Kahn S., Grindlay J., 1983a// Astrophys. J.-V.266.-P.160.

Leahy D., Elsner R., Weisskopf M., 1983b// Astrophys. J.-V.-272.-P.256.

Le Gland F., 1982// Techn. Interf. Tres. Grande Base. Coll. Int. Toulouse.- Toulouse.-P.303.

Leotta - Janin C., 1967// Bull. Astr. Inst. Neth.- V.19.-P.169.

Levinson H., 1947// J. Math. Phys.- V.25.-P.261.

Liversey A.K., Skilling J., 1984// Acta Crystallogr., A.

Lomb N.R., 1976// Astrophys. Sp. Sci.- V.39.-P.447.

Loudon R., 1973; The Quantum Theory of Light.- Oxford: Clarendon Press; перевод: Лоудон Р., 1976; Квантовая теория света.- М.: Мир.

Loumos G.L., Deeming T.J., 1978// Astrophys. Sp. Sci.-V.56.-P.285.

Lucy B., 1976// Astrophys. J.- V.206.-P.499.

Lynden-Bell D., 1969// Nature.- V.223.-P.690.

Mandel L., 1958// Proc. Phys. Soc. (London).- V.72.-P.1037.

Mandel L., 1959// Proc. Phys. Soc. (London).- V.74.-P.233.

Mandelbrot B.B., 1977; Fractals: Form, Chance and Dimension.- San Francisco: Freeman.

Manwell T., Simon M., 1966// Nature.- V.212.-P.1224.
Manwell T., Simon M., 1968// Astron. J.- V.73.-P.407.
Mardia K.V., 1972; Statistics of Directional Data.- London; New York: Academic Press; перевод: Мардиа К., 1978; Статистический анализ угловых наблюдений.- М.: Наука.
Marraco H.G., Muzzio J.C., 1980// Publ. Astr. Soc. Pacific. -V.92.-P.700.
Marsh K.A., Richardson J.M., 1987// Astron. Astrophys.- V.182.-P.174.
Marvasti F., 1986// Proc. IEEE.- V.74.-P.1596.
Matthews M.V., Pierce J.R., 1987// Scientific American.- V.256, No. 2.
Mc Hardy I., Czerny B., 1987// Nature.- V.325.-P.696.
Meekins J.F., Wood K.S., Hedler R.L., Byram E.T., Yentis D.J., Chubb T.A., Friedman H., 1984// Astrophys. J.- V.278.-P.288.
Mehta C.L., 1970// Progress in Optics/ Ed. E. Wolf.- Amsterdam, London: North-Holland.- V.8.-P.375.
Meisel D.D., 1978// Astron. J.- V.83.-P.538.
Meisel D.D., 1979// Astron. J.- V.84.-P.116.
Moles M., Garcia-Pelayo J.M., Masegosa J., Garrido R., 1986//Astron. J.- V.92.-P.1030.
Morbey C.L., 1978a// Publ. Dominion Astrophys. Obs.- V.14.-P.185.
Morbey C.L., 1978b// Publ. Dominion. Astrophys. Obs.- V.15.-P.105.
Mushotzky R.F., 1987// Nature.- V.325.-P.666.
Nemec A.F., Nemec J.M., 1985// Astron. J.- V.90.-P.2317.
Nutall A.H., Carter G.C., 1982// Proc. IEEE.- V.70.-P.1115.
Nyquist H., 1924// Bell Syst. Techn. J.- No.4.-P.324.
Nyquist H., 1928// AIEE Trans.- V.47.-P.617.
Pacholczyk A.G., Lubart N.D., Penning W.R., Ferguson D.H., Turnshek D., 1983// Astrophys. Lett.- V.23.-P.225.
Papoulis A., 1985// Proc. IEEE.- V.73.-P.1332.
Parzen E., 1961// Technometrics.- V.3.-P.167.
Pearson K., 1900// Phil. Mag.-V.50.-P.157.
Pelt J., 1983// Statistical Methods in Astronomy/ Ed. E.J. Rolfe - Noordwick: ESA Scient. and Techn. Publ.-P.37.

Penston M.V., Penston M.J., Neugebauer G., Tritton K.P., Becklin E.E., Visvanathan N., 1971// Monthly Notices Roy. Astron. Soc.- V.153.-P.29.

Penston M.V., Penston M.J., Selmes R.A., Becklin E.E., Neugebauer G., 1974// Monthly Notices Roy. Astron. Soc.- V.169.-P.357.

Percy J.R., 1977// Monthly Notices Roy. Astron. Soc.- V.181.-P.647.

Percy J.R., Jakate S.M., Mathews J.M., 1981// Astron. J.- V.86.-P.53.

Poincare H., 1912// Chance Monist.- V.22.-P.31.

Ponman T., 1981// Monthly Notices Roy. Astron. Soc.- V.196.-P.583.

Press W.H., 1978// Comments Astrophys.- V.7.-P.103.

Press W.H., Flannery B.P., Teukolsky S.A., Vetterling W.T., 1986; Numerical Recipes.The Art of Scientific Computing. - Cambridge Univ. Press.

Press W.H., Schechter P., 1974// Astrophys. J.- V.193.- P.437.

Rabiner L.R., Gold B., 1975; Theory and Application of Digital Signal Processing.- New York: Prentice-Hall; перевод: Рабинер Л., Голд Б., 1978; Теория и применение цифровой обработки сигналов.- М.: Мир.

Rayleigh, Lord. (J.W. Strutt), 1880// Phil. Mag.- V.10.-P.73.

Reed I.S., 1962// IRE Trans.-IT-8, No.3.-P.194.

Rees M.J., 1978// Physica Scripta.- V.17.-P.193.

Renson P., 1978// Astron. Astrophys.- V.63.-P.125.

Richter H.B., Ulrych T.J., 1974// Astrophys. J.- V.192.-P.719.

Roberts D.H., Lehar J., Dreher J.W., 1987// Astrophys. J.- V.93.-P.968.

Robinson E.A., 1982// Proc. IEEE.- V.70.-P.885.

Saletti R., 1986// Proc. IEEE.- V.74.-P.1595.

Salpeter E.E., 1964// Astrophys. J.- V.140.-P.796.

Sandage A.R., 1965// Astrophys. J.- V.141.-P.1560.

Scargle J.D., 1981// Astrophys. J., Suppl.- V.45.-P.1.

Scargle J.D., 1982// Astrophys. J.- V.263.-P.835.

Schmidt M., 1963// Nature.- V.197.-P.1040.
Schuster A., 1894// Phil. Mag.- V.37.-P.509.
Schuster A., 1897// Proc. Roy. Soc.-V.61.-P.455.
Schuster A., 1898// Terrest. Magn.- V.3.-P.13.
Schuster A., 1904// The Theory of Optics.- London: Cambridge Univ. Press.
Schuster A., 1906a// Proc. Roy. Soc.- V.77.-P.136.
Schuster A., 1906b// Phil. Tr. Roy. Soc., Ser. A.- V.206.-P.69.
Schwarz U., 1978// Astron. Astrophys.- V.65.-P.345.
Seyfert C.K., 1943// Astrophys. J.- V.97.-P.28.
Shannon C., 1948// Bell Syst. Techn. J.- V.27.-P.379,623.
Shannon C., 1949// Proc. IRE.- V.37.-P.10.
Shapley H., 1914// Astron. Nachr.-Bd.196.-S.418.
Shore J.E., *Johnson R.W.*, 1980// IEEE Trans.- IT-26.-P.26.
Shore J.E., *Johnson R.W.*, 1983// IEEE Trans.- IT-29.-P.942.
Simpson G., *Mayer-Hasselwander H.*, 1986// Astron. Astrophys.- V.162.-P.340.
Slipher V.M., 1917// Lowell Obs. Bull.- V.3.-P.59.
Smith H.J., 1965// Quasi-Stellar Sources and Gravitational Collapse/ Eds. I.Robinson, A. Schild, E.L. Shucking.- Chicago: Univ. of Chicago Press.
Smith H.J., *Hoffleit D.*, 1963// Nature.- V.198.-P.650.
Snyder D.L., 1975// Random Point Processes.- New York: Wiley.
Stellingwerf R.F., 1978// Astrophys. J.- V.224.-P.953.
Stobie R.S., 1977//Monthly Notices Roy. Astron. Soc.- V.180.-P.631.
Stralkowski C.M., 1968// Phil. D. Thesis. Univ. of Wisconsin.
Strong A.W., 1985// Astron. Astrophys.- V.150.-P.273.
Sutherland P.G., *Weisskopf M.C.*, *Kahn S.M.*, 1978// Astrophys. J.- V.219.-P.1029.
Swan P.R., 1982// Astron. J.- V.87.-P.1608.
Swingler D.N., 1985// Astron. J.- V.90.-P.675.
Tanner R.W., 1948// Comm. David Dunlap. Obs.- V.16.-P.177.
Terrel J., *Olsen K.H.*, 1970// Astrophys. J.- V.161.-P.399.
Thompson A.R., *Moran J.M.*, *Swenson G.W.*, *Jr.*, 1986;

Interferometry and Synthesis in Radio Astronomy.- New York: Wiley.

Thomson D.J., 1982// Proc. IEEE.- V.70.-P.1055.

Timothy J.G., 1988// Instrumentation for Ground-Based Optical Astronomy/ Ed. L.B. Robinson.- Berlin. Springer--Verlag.-P.516.

Ulrych T.J., 1972a// J. Geoph. Res.- V.77.-P.1396.

Ulrych T.J., 1972b// Nature.- V.235.-P.218.

Ulrych T.J., Bishop T.N., 1975// Rev. Geophys. and Sp. Phys.-V. 13:1.-P.183.

Ulrych T.J., Clayton R.W., 1976// Phys. of the Earth and Planet.Inter.- V.12.-P.188.

Urry C.M., 1988// Space Tel. Sci. Inst.Prepr.Ser.-No.270.- P.1.

Van den Bos A., 1971// IEEE Trans. Inform. Theory.-IT-17. - P.493.

Van der Ziel A., 1976// Noise in Measurements.- New York: Wiley; перевод: Ван дер Зил А., 1979; Шумы при измерениях.- М.: Мир.

Vanicek P., 1971// Astrophys. Sp. Sci.- V.12.-P.10.

von Neumann J., 1941// Ann. Math. Stat.- V.12.-P.367.

von Neumann J., 1942// Ann. Math. Stat.- V.13.-P.86.

von Neumann J., Kent R.H., Bellinson H.R., Hart B.I., 1941// Ann. Math. Stat.- V.12.-P.153.

Walker G., 1914// Calcutta Ind. Met. Mems.- V.21.-P.9.

Walker G., 1931// Proc. Royal Soc., Ser. A.- V.131.-P.518.

Warner B., Robinson E.L., 1972// Monthly Notices Roy. Astron. Soc.-V.159.-P.101.

Webb J.H., 1972// Am. Journ. Phys.- V.40.-P.850.

Wehlau W., Leung K.-C., 1964// Astrophys. J.- V.139.-P.843.

Weisskopf M.C., Kahn S.M., Sutherland P.G., 1975// Astrophys. J. Lett.-V.199, L147.

Whittaker E., Robinson G., 1924; The Calculus of Observations.-London: Blackies.

Whittaker J., 1935; Interpolatory Function Theory.- Cambridge Tracts in Math. and Math.Phys., No.33.- Cambridge Univ. Press, IV.

Whitworth W.A., 1959 (5-th ed. 1901); Choice and chance.-

Cambridge: Deighton Bell.

Wiener N., 1933; The Fourier Integral and Certain of Its Applications.- New York: Cambridge Univ. Press.

Wiener N., 1949; Extrapolation, Interpolation and Smoothing of Stationary Time Series.- New York: Wiley.

Wold H., 1965; Bibliography on Time Series and Stochastic Processes.- Cambridge: MIT Press.

Yule G.U., 1927// Phul. Trans. Roy. Soc. London, Ser. A.- V.226.-P.267.

Zamorani G., Giommi P., Maccacaro T., Tananbaum H., 1984// Astrophys. J.- V.278.-P.28.

УКАЗАТЕЛЬ ИНОСТРАННЫХ АВТОРОВ

Акаике (Akaike H.)
Аллен (Allen C.W.)
Андерсен (Andersen N.)
Андерсон Р. (Anderson R.L.)
Андерсон Т.(Anderson T.W.)
Андреасен (Andreasen G.K.)
Андропов (Andropov I.L.)
Балиунас (Balčiunas S.L.)
Барбиери (Barbieri L.)
Барнинг (Barning F.J.M.)
Бартлетт (Bartlett M.S.)
Баруча–Рнд (Bharucha–Reid A.T.)
Берг (Burg J.P.)
Берри (Berry P.A.M.)
Берримен (Berryman J.G.)
Бертолотти (Bertolotti M.)
Бишоп (Bishop T.N.)
Блумфилд (Bloomfield P.)
Блэкмен (Blaikman R.B.)
Бой (Boy W.R.)
Бойнтон (Boynton P.E.)
Бокс (Box G.E.P.)
Боксенберг (Boksenberg A.)
Болдт (Boldt E.A.)
Бопп (Bopp B.W.)
Борн (Born M.)
Браун (Brown R.H.)
Брейсуэлл (Bracewell R.N.)
Бриллинджер (Brillinger D.R.)
Бролт (Brault J.W.)
Брэдли (Bradley J.U.)
Буре (Bourret R.C.)
Буччери (Buccheri R.)
Бэйли (Bailey S.I.)
Бэрк Г. (Burki G.)
Бэрк Э. (Burke E.W.)
Вайскопф (Weisskopf M.C.)
Ван ден Бос (Van den Bos A.)
Ван дер Зил (Van der Ziel A.)
Ван Винкль (Van Winkle W.)
Ванишек (Vanicek P.)
ван Хаутен (van Houten)
Ваттс (Watts D.G.)
Вейлау (Wehlau W.)
Винер (Wiener N.)
Волд (Wold H.)
Вольф (Wolf E.)
Гибсон (Gibson A.I.)
Голд (Gold B.)
Голдман (Goldman S.)
Грей (Gray D.F.)
Грот (Groth E.J.)
Гэлл (Gull S.F.)
Давенпорт (Davenport W.B.)
Дворецкий (Dworetsky M.M.)
Дезикачарн (Desikachary K.)
Делаш (Delache P.)
Джакате (Jakate S.M.)
Джейнс (Jaynes E.T.)

Дженкинс (Jenkins G.M.)
Джентлмен (Gentleman W.M.)
Джонсон (Johnson R.W.)
Джоунс (Jones C.)
Джуди (Jeudy L.M.A.)
Диаконис (Diaconis P.)
Диминг (Deeming T.J.)
Дитер (Deeter J.E.)
Дрехер (Dreher J.W.)
Дэвис (Davis H.T.)
Дюваль (Duvall T.L.)
Дюмон (Dumont Th.)
Дю Пюи (Du Puy D.L.)
Дютта (Dutta P.)
Заморани (Zamorani G.)
Иэди (Eadie W.T.)
Камминс (Cummins H.Z.)
Канасевич (Kanasewich E.R.)
Кано (Kano Y.)
Картер (Carter G.C.)
Кендалл (Kendall M.)
Кинман (Kinman T.D.)
Кирк (Kirk B.L.)
Клейтон (Clayton R.W.)
Ковач (Kovacs G.)
Кокс (Cox D.R.)
Колгейт (Colgate S.A.)
Коллура (Collura A.)
Крамер (Kramer H.)
Кроу (Crow E.L.)
Крэддок (Craddock J.M.)
Кули (Cooley J.W.)
Кун (Kuhn J.R.)
Кункель (Kunkel W.E.)
Курри (Currie R.G.)
Курц (Kurtz D.W.)
Кэмпбелл (Campbell N.)
Лагранж (Lagrange J.L.)
Лакосс (Lacoss R.T.)
Ланцош (Lanczos C.)
Лафлер (Lafler J.)
Левинсон (Levinson H.)
Ле Глан (Le Gland F.)
Леотта-Жанин (Leotta-Janin C.)
Леунг (Leung K.-C.)
Ливерси (Liversey A.K.)
Линден-Белл (Lynden-Bell D.)
Лихи (Leahy D.A.)
Ломб (Lomb N.R.)
Лоудон (Loudon R.)
Лоуренс (Lawrence A.)
Лумос (Loumos G.L.)
Льюис (Lewis P.A.W.)
Лэмптон (Lampton M.)
Люси (Lusy B.)
Мак Харди (Mc Hardy I.)
Мандел (Mandel L.)
Манделброт (Mandelbrot B.B.)
Марвасти (Marvasti F.)
Мардиа (Mardia K.V.)
Маррако (Marraco H.G.)
Машоцки (Mushotzky R.F.)
Мейер-Хассельвандер (Mayer-Hasselwander H.)
Мейсел (Meisel D.D.)
Мета (Mehta C.L.)
Микинс (Meekins J.F.)
Моле (Moles M.)
Морби (Morbey C.L.)
Муццио (Muzzio J.C.)
Мэдер (Maeder A.)
Мэнвелл (Manwell T.)
Мэттьюз (Mattthews J.M.)
Мэттьюз (Matthews M.V.)
Мэрш (Marsh K.A.)
Найквист (Nyquist H.)
Немек А. (Nemec A.F.)
Немек Дж. (Nemec J.M.)
Нэталл (Nutall A.H.)

Олсен (Olsen K.H.)
Пайк (Pike E.R.)
Папулис (Papoulis A.)
Парзен (Parzen E.)
Пауэлл (Powell M.J.D.)
Пахольчик (Pacholczyk A.G.)
Пельт (Pelt J.)
Пенстон (Penston M.V.)
Перси (Percy J.R.)
Петерсен (Petersen J.O.)
Пирс (Pierce J.R.)
Пирсон (Pearson K.)
Понман (Ponman T.)
Пресс (Press W.H.)
Пуанкаре (Poincare H.)
Рабинер (Rabiner L.R.)
Раст (Rust B.W.)
Рафинер (Rufener F.)
Ренсон (Renson P.)
Рид (Reed I.S.)
Рис (Rees M.J.)
Рихтер (Richter H.B.)
Ричардсон (Richardson J.M.)
Робертс (Roberts D.H.)
Робинсон Г. (Robinson G.)
Робинсон Э. (Robinson E.L.)
Ролланд (Rolland W.W.)
Рут (Root W.L.)
Рэйн (Raine D.)
Рэлей (Rayleigh)
Рэлстон (Ralston A.)
Саймон (Simon M.)
Салетти (Saletti R.)
Сван (Swan P.R.)
Свинглер (Swingler D.N.)
Сейферт (Seyfert C.K.)
Симпсон (Simpson G.)
Скаргл (Scargle J.D.)
Скиллинг (Skilling J.)
Слайфер (Slipher V.M.)
Смит (Smith H.J.)
Снайдер (Snyder D.L.)
Солпитер (Salpeter E.E.)
Стеллингвеф (Stellingwerf R.F.)
Стоби (Stobie R.S.)
Стралковски (Stralkowski C.M.)
Стронг (Strong A.W.)
Стюарт (Stuart A.)
Сэзерланд (Sutherland P.G.)
Сэндидж (Sandage A.R.)
Тананбаум (Tananbaum H.)
Твисс (Twiss R.Q.)
Террел (Terrel J.)
Тимоти (Timothy J.G.)
Томсон А. (Thompson A.R.)
Томсон Д. (Tomson D.J.)
Тьюки (Tukey J.W.)
Тьюколски (Teukolsky S.A.)
Тэннер (Tanner R.W.)
Уайт (White O.R.)
Уайтворт (Whitworth W.A.)
Уилф (Wilf H.S.)
Уиттекер Дж. (Whittaker J.)
Уиттекер Э. (Whittaker E.)
Ульрих (Ulrych T.J.)
Уокер (Walker G.)
Уорнер (Warner B.)
Урри (Urry C.M.)
Уэбб (Webb J.H.)
Фейнман (Feynman R.P.)
Фелжетт (Fellgett P.)
Феллер (Feller W.)
Фернн (Fernie J.D.)
Феррас–Мелло (Ferraz–Mello S.)
Филд (Field G.B.)
Фитч (Fitch W.S.)
Фишер (Fisher R.A.)
Фланнери (Flannery B.P.)
Флетчер (Fletcher R.)
Фолкнер (Faulkner D.J.)

фон Нейман (von Neumann J.)
Формэн (Forman W.)
Фраска (Fraska S.)
Фриден (Frieden B.R.)
Фуже (Fourgere P.F.)
Фуллер (Fuller W.)
Фэлмен (Fahlman G.G.)
Фэт (Fath E.A.)
Хаген (Hagen J.G.)
Хадсон (Hudson D.J.)
Харви (Harvey J.W.)
Харт (Hart B.I.)
Хек (Heck A.)
Хемминг (Hamming R.W.)
Хеннан (Hannan E.J.)
Хеннинг (Henning H.M.)
Хегбом (Hogbom J.A.)
Хиллс (Hills J.G.)
Хорман (Horman M.)
Хорн Дж. (Horne J.H.)
Хорн П. (Horn P.M.)
Хуг (Hooge F.N.)
Хоффлейт (Hoffleit D.)
Хоффман (Hoffman G.A.)
Хоффмейстер (Hoffmeister C.)
Чандрасекар (Chandrasekhar S.)
Черны (Czerny B.)
Чэтфилд (Chatfield C.)
Шварц (Schwarz U.)
Шеннон (Shannon C.)
Шепли (Shapley H.)
Шеррер (Scherrer P.H.)
Шехтер (Schechter P.)
Шмидт (Schmidt M.)
Шор (Shore J.E.)
Шустер (Schuster A.)
Эберт (Ebert K.)
Эдерер (Ederer H.)
Эйблс (Ables J.G.)
Эйри (Airy G.B.)
Эльст (Elst E.W.)
Энжион (Angione R.J.)
Энслейн (Enslein K.)
Эренфест П. (Ehrenfest P.)
Эренфест Т. (Ehrenfest T.)
Эфрон (Efron B.)
Юл (Yule G.U.)
Юркевич (Jurkevich I.)

Аподизация 103
Боковые лепестки 108
Временной ряд 10, 11, 76
Гармоническое колебание 38
Год (сидерический, тропический) 279
Дирака
- гребень 46
- функция 357
Дисперсионное отношение 262
Дисперсионный анализ 260
Дифракционная решетка 232
Духи
- Лаймана 232
- Роуланда 232
Закон арксинуса 221
Звезда
- *BK* Cen 200
- *CF* Cas 66
Значимость
- критерий 89
- максимумов периодограммы 96
- тренда 215
- уровень 89
Инварианты фильтра 319, 340
Индекс дисперсии 290
Интенсивность 333, 336
Информация 116
Квазары 149
Квантовая эффективность обнаружения 348
Квантовый выход 331
Когерентность 166
Коспектр 164
Коэффициент
- вариации 214
- диффузии 299
Критерий
- случайности 201
- χ^2 313
Ломба представление 175
Лоренца профиль 36, 335
Метод
- максимума энтропии 110, 111
- сворачивания (наложения) эпох 72
Модель
- авторегрессионная 125, 128
- параметрическая 127
Обратная задача 29, 243
Окно Бартлетта 108
- временное 62
- корреляционное 105
- прямоугольное 108
- спектральное 63
- - равномерного ряда 77

Окончательная ошибка предсказания Акаике 124
Отбеливание 19, 186
- гармоническое 187
- неравномерных рядов 191
- полигармоническое 188
Оценка
- максимального правдоподобия 171
- несмещенная 101
- параметров авторегрессии 140
- состоятельная 60, 101
- Юла–Уокера 141
Периодограмма
- Аббе–Лафлера–Кинмана 71, 257
- взаимная 168
- гармонического колебания 79
- случайного ряда 90
- Шустера 15
- - временного ряда со случайными моментами наблюдения 180
Плотность распределения вероятностей 362
Подмена частот (aliasing) 16, 17, 83, 274
Приближение
- Фраунгофера 240
- Френеля 239
Разрешение 41, 69
Распределение
- биномиальное 288, 362
- гамма 352
- Гаусса (нормальное) 33, 364
- геометрическое 114, 339, 363
- Кокса–Мандела 297
- прямоугольное 33, 363
- Пуассона 288, 362
- Рэлея 95
- Уокера 97
- Фишера 98
- χ^2 365
- экспоненциальное 364
- Эрланга 352
Свертка 31
Сглаживание
- временного окна 102
- корреляционного окна 104
Сейфертовские галактики 148, 216
Система
- нормальных уравнений 172
- уравнений Юла – Уокера 141
Скважность 45
Случайная величина 360
Случайный процесс (случайная функция) 46
- - авторегрессионный 128
- - - второго порядка (ряд Юла) 135
- - - первого порядка 132
- - блуждания (броуновского движения) 219
- - восстановления 300
- - марковский 132
- - периодический стационарный 51
- - Пуассона 285

– – – гармонический 295, 321
– – – дважды стохастический 296, 338
– – – неоднородный 291
– – – однородный 286
– – – ступенчатый 294, 326
– – скользящего среднего 129
– – стационарный 47
– – точечный 284, 301
Спектр
– взаимный амплитудный, фазовый 164
– дискретный 62
– – равномерного ряда 79
– импульсный 307
– квадратурный 164
– *LS* 173, 174
– LS_m 179
– мощности 42
– функции 31
– энергетический 33
Спектральная плотность 43, 48
– – авторегрессионного процесса 131
– – – – второго порядка 140
– – взаимная 163
– – выборочная 91
– – интенсивности излучения 334
– – марковского процесса 134
– – электромагнитных колебаний 333
Скользящее среднее 213
Статистика
– Аббе (– Лафлера – Кинмана) 216, 256
– D_2 179
– *LS*–спектра 177
– Рэлея 307
– Фишера 196
Сутки (звездные, солнечные) 278
Счетчики I и II рода 345
Теорема
– Винера – Хинчина 49
– отсчетов (Котельникова – Шеннона) 85
– Парсеваля 81
– Рэлея 32, 44, 49
Тихонова регуляризующий алгоритм 251
Тренд 12, 208
– линейный 208
– полиномиальный 211
Узкополосный сигнал 334
Уравнение
– характернстическое 135
– типа свертки 243
Условия стационарности авторегрессионного процесса 138
Утечка энергии 108
Фаза 18, 38
Фазовая диаграмма 18
Фильтр
– гомоморфный 250
– инверсный 248
– линнейный 53, 163, 165, 237
– оптимальный Винера 248
– предсказания 125
– управляемый 250

Фликкер–шум 57, 218
Формула
– Барнинга 174
– Берримена 124
– Кемпбелла 151
– Шеннона 118
– Эйри 242
Фрактал 218
Фрактальная размерность 231
Функция
– автоковариационная 37, 43, 47
– – выборочная 84
– – интенсивности излучения 334
– взаимная ковариационная 162
– Грина 55
– передаточная 55, 238
– переходная 31
– периодическая 37
– почти–периодическая 39
– производящая 361
– – распределения Пуассона 289
– рассеяния точки (аппаратная функция) 55, 238, 243
– случайная, см. случайный процесс
– с ограниченным спектром 85
– ступенчатая 45
– характеристическая 182
Фурье интеграл 31
– коэффициенты 39
– ряд 38
– преобразование 31
Характеристика
– импульсная 54, 237
– фазо–контрастная 238
– частотно–контрастная 238
Частота
– Найквиста 16
– отсчетов 87
– сопряженная (ложная, подмены) 16, 24, 274–278
– среза 249
Число Вольфа 142
Шум
– белый 49
– – гауссовский 50
– дробовой 150
– пуассоновскнй 297
Энтропия
– дифференциальная 114
– – стационарного случайного процесса 117
– совместная 115
– средняя условная 116
Эргодичность 56
Ядра галактик 148, 216
Ядро Фейера 109, 360

Научное издание
ТЕРЕБИЖ Валерий Юзефович
Анализ временных рядов в астрофизике

Заведующий редакцией Н.А. Носова
Редактор О.В. Салецкая
Художественный редактор Г.М. Коровина

Операторы: Г.Г. Короткова, Л.И. Спангенберг
Корректоры: Т.Г.Егорова, Л.С. Сомова

ИБ № 41132
Компьютерный набор. Подписано к печати 23.12.91.
Бумага тип. № 1
Формат 60x90/16. Печать офсетная.
Усл. печ. л. 24,5 Усл. кр.-отт. 24,5 Уч.-изд. л. 24,28.
Тираж 720 экз. Заказ № 1146

Издательско-производственное и книготорговое объединение
"Наука"
Главная редакция физико-математической литературы
117071 Москва В-71, Ленинский проспект, 15

Отпечатано в Четвертой типографии издательства "Наука"
630077 г.Новосибирск 77, Станиславского, 25

Valery Yu.Terebizh

TIME SERIES ANALYSIS IN ASTROPHYSICS
Moscow, Nauka, Main Editorial Board for Physical and Mathematical Literature, 1992

Readership: Astronomers, physicists, specialists in applied mathematics, postgraduates and students. It can also be useful for researchers in other sciences (biology, engineering) who deal with study of the intrinsic structure of time processes.

The book: The book conserns the analysis of time series, in particular, searching for periodicities in data sampled at irregular observational intervals and the counting of events. Equally with the classical approach to the estimating of power spectrum the modern methods are considered as well (MAXIMUM ENTROPY method, CLEAN algorithm etc.). The considerable part of contents for the first time is given in a monographic aspect, some results are new. The concept of time series as an ordered sample of measurements includes a very extensive range of observations in physics, economy and techniques. For known methods as well as for the new ones the computational algorithms are given. General propositions are illustrated by numerous examples, which in every case include the initial data, formulae for computing and the results of analisis in the graphical form. The astronomical context of the majority of examples does not restrict the applicability of methods to the problems of geophysics, biology, economics and other branches of science, where periodicity, power spectrum and the intrinsic model of time variability should be investigated.

Contents: **Fundamental concepts**. An introducting example. General discussion of the problem. Fourier integral. Periodic functions. Spectral analysis of stationary stochastic functions. Real processes. Two approaches to the search for periodicity. **Equidistant time series.** Main equations Sampling theorem. Statistical properties of the periodogram. Smoothing of the periodogram. Spectrum estimation with the aid of Maximum Entropy Method. Autoregression models Variations of the Solar activity. Quasar 3C 273 light curve analysis. Mutual power spectrum of two processes. **Spectral theory of non-equidistant time series.** LS_m-spectrum. Randomly spaced observations. Whitening of time series The CLEAN algorithm. Trend detection. Flicker-noise. Flux variability of Seyfert nuclei. Optical analogies. General approach to the solving of inverse problems. **Non-parametric research for possible periodicity.** Abbe Lafler – Kinman statistics. Phase binning of data points (Whittaker – G.Robinson's, Warner E.Robinson's and Stellingwerf's methods). Methods that retain efficiency for sparse measurements ("string length", Renson's and Pelt's) Autocorrelation methods The aliasing problem. A comparative study of different periodfinding methods. **Pulse random sequences.** Poisson process. Spectral analysis of pulse sequences Methods based on binning of events. Examples. The photon counting. Dead-time effects. Signal to noise ratio for events counting. **Afterword**. **Appendices.** Useful sums. Main formulae with the Dirac's function. Characteristics of random variables. Percentage points for the ratio of largest periodogram ordinate to the average value. Percentage points for the Abbe criterion Numerical data. **List of literature**

The author: Valery Yu. Terebizh, D.Sc. (Phys. & Math.), is head of the Creamean Station of Sternberg Astronomical Institute of Moscow University. His field of interest is the physics of stars and galaxies, and treatment of results of astrophysical observations.