У.Гренандер

ВЕРОЯТНОСТИ НА АЛГЕБРАИЧЕСКИХ СТРУКТУРАХ

Книга известного шведского математика У. Гренандера «Вероятности на алгебраических структурах» содержит изложение современных разделов теории вероятностей, развитых в самые последние годы. В ней отчетливо отражены связи теории вероятностей с другими разделами современной математики, особенно с алгеброй и топологией.

Книга представляет большой интерес не только для тех, кто занимается теорией вероятностей, но и для математиков других специальностей, а также для физиков, научных работников и инженеров, использующих в своих исследованиях методы и приложения теории вероятностей.

ОГЛАВЛЕНИЕ

УКАЗАТЕЛЬ

ПРЕДИСЛОВИЕ РЕДАКТОРА ПЕРЕВОДА

Книга известного шведского математика У. Гренандера в настоящее время является единственной в своем роде. Она посвящена предельным теоремам на алгебраических структурах, в первую очередь на группах. Под предельными теоремами здесь понимаются теоремы о поведении распределений сумм большого числа независимых случайных элементов со значениями в группах и т. п. С аналитической точки зрения дело сводится к изучению композиций (свертки) большого числа распределений на группе. Анализ Фурье (базирующийся на теории характеров или теории представлений) является основным орудием.

В первой главе автор убедительно доказывает желательность перенесения результатов классической теории предельных теорем на случай групп и других алгебраических структур, и поэтому нет необходимости говорить об этом здесь.

Уместно напомнить, что классическая теория рассматривает или суммы

$$\zeta_n = \xi_{n,1} + \xi_{n,2} + \ldots + \xi_{n,k_n}$$

бесконечно малых независимых случайных величин («схема серий»), или суммы

$$\zeta_n = \xi_1 + \xi_2 + \ldots + \xi_n,$$

образованные по последовательности независимых (чаще всего одинаково распределенных) случайных величин. «Схема серий» (тесно связанная с теорией случайных процессов с независимыми приращениями) представляет собой достаточно трудную проблему для изучения на группах. Даже на мультипликативной группе не равных нулю комплексных чисел можно, как показал В. М. Золотарев, построить весьма содержательную теорию.

Схема последовательности (ограничимся для простоты одинаково распределенными величинами) поддается полному изучению в случае компактных групп, где предельные распределения, по существу, исчерпываются мерой Хаара на группе и ее подгруппах. Однако в очень интересном случае локально компактных некоммутативных групп в момент написания этой книги даже сама задача не была точно поставлена. Распределение сумм s_n «расплывается» на группе, и в то время как на числовой прямой это «расплывание» может быть компенсировано нормировкой, т. е. переходом от s_n к $(s_n - a_n)/b_n$ (a_n и b_n — постоянные, $b_n \to \infty$), в общих локально компактных группах это свойство отсутствует. Совсем недавно в работах, в частности Тутубалина и Ферстенберга, был поставлен и решен в отдельных случаях вопрос о том, когда предельное поведение распределения s_n определяется конечным числом «параметров», т. е. когда существует конечное число функционалов f_k, таких, что из $f_k(P) = f_k(Q)$ следует

$$P^{n*}(A) - Q^{n*}(A) \to 0$$

для достаточно широкого класса подмножеств A рассматриваемой группы.

Книга написана не очень ровно, и внимательный читатель сам заметит ее слабые места. Но она будет весьма полезной математикам различных специальностей особенно потому, что в ней показано взаимодействие таких разделов математики как алгебра, теория вероятностей, топология и функциональный анализ. Читатель увидит, как много интересных нерешенных проблем остается в увлекательной области, которой посвятил свою книгу автор.

Ю. Прохоров

Моим родителям

ПРЕДИСЛОВИЕ АВТОРА

В этой книге делается попытка представить единую и последовательную теорию исчисления вероятностей на алгебраических структурах. Среди таких структур мы рассмотрим некоторые, заслуживающие особого внимания ввиду роли, которую они играют в приложениях этой теории, а также их внутреннего интереса. Это топологические полугруппы и группы, топологические векторные пространства и алгебры.

У автора интерес к этой теме возник в связи с некоторыми практическими задачами, часть которых описана в главе 1. Поэтому в книге особое внимание уделяется конкретным результатам. Эти результаты могут быть фактически использованы, и они приводят, по крайней мере в принципе, к решениям, которые можно определить точно или приближенно, аналитически или численно. Мы не пытались добиться результатов наиболее общего характера или развить эту теорию в наиболее совершенной форме. Поэтому мы, не колеблясь, налагали условия, подобные сепарабельности, измеримости по Борелю и т. д., когда они приводили к упрощениям, хотя и не всегда были необходимыми. Некоторые читатели могут почувствовать склонность к устранению подобных недочетов, к получению необходимых и достаточных условий и т. д. Автор надеется, что в будущем это кому-нибудь удастся сделать.

Чтобы избежать затемнения главных идей теории длинными выкладками и рассуждениями технического характера, такие выкладки иногда только лишь намечаются. Это особенно относится к случаям, когда можно отослать читателя к литературе, в которой дается полное изложение вопроса.

Для того чтобы сделать изложение как можно более конкретным, в конце каждой главы дано несколько специальных примеров. Автор полагает, что некоторые из этих примеров не менее важны, чем теоремы, которые они иллюстрируют. Иногда они могут указывать возможные пути развития теории.

Читатель должен заметить, что все ссылки, а также дополнительная информация даются в замечаниях в конце книги. Глава I содержит исторический очерк и некоторые замечания относительно практических основ теории. Рассуждения в этой главе имеют эвристический характер и повторяются в строгой форме в последующих главах. При первом чтении читатель может пропустить разд. 4.2 (где только намечаются доказательства, имеющие технический характер), 5.4 и 6.6 (относящиеся к специальным вопросам).

Трудно точно определить необходимые для чтения книги знания, но представляется ясным, что стандартный курс «высшей математики» вряд ли будет достаточен. Так как эта книга написана для вероятностников, то предполагается, что читатель хорошо знаком с теорией вероятностей и теорией меры, скажем, в объеме, соответствующем содержанию книг Лоэва «Теория вероятностей» и Халмоша «Теория меры». Многие из используемых рассуждений, относящихся к теории меры, имеют стандартный характер и только намечены. Читатель должен также обладать некоторым представлением об основных понятиях топологической алгебры. Горячо рекомендуется книга Наймарка «Нормированные кольца», дающая ясное и современное изложение вопроса. Так как эта область, быть может, не так хорошо известна специалистам по теории вероятностей, то представляется целесообразным более подробное обсуждение соответствующих вопросов. В замечаниях читатель найдет некоторое число основных определений и логических связей. Мы также требуем от читателя знания элементов функционального анализа. Подходящим руководством может служить книга Хилле и Филлипса «Функциональный анализ и полугруппы», содержащая также весьма полезные сведения о полугруппах, почти необходимые при изучении однородных процессов.

Объектом нашего исследования является распределение вероятностей на структуре, на которой определена по меньшей мере одна бинарная алгебраическая операция, непрерывная в соответствующей топологии. В этом случае можно говорить о стохастическом элементе, извлеченном случайно из этой структуры. Соединяя два таких независимых стохастических элемента бинарной операцией, мы приходим к композиции (вот ключевое слово!) двух распределений вероятностей. Изучение поведения композиций, особенно в случае их большого числа, будет занимать значительное место. Наш подход будет аналитическим, основным инструментом будет анализ Фурье. Нет сомнения в правильности такого подхода, если мы хотим получить определенные результаты, алгоритмы и т. д. Существует, однако, и другой путь, более алгебраический или, быть может, даже вероятностный по форме. Именно, мы рассматриваем множество распределений вероятностей, о которых идет речь, как топологическую полугруппу. Применяя теорию полугрупп, можно получить очень интересные результаты. Этот подход действительно является довольно привлекательным применением общей теории полугрупп, но он только иногда будет обсуждаться в тексте.

Я имел несколько ценных обсуждений с коллегами, которых мне хотелось бы поблагодарить за их советы: Р. Форте, Е. Мурье, Л. Шметтерер, З. Шидак и А. Шпачек. Я очень признателен Д. Вену за предоставление мне возможности ознакомиться с его рукописью до ее опубликования. Разд. 4.2—4.4 существенно опираются на результаты Вена и его изложение. М. Розенблат и В. Фрейбейгер прочли всю рукопись моей книги и предложили многие изменения. Р. Лойнес и П. Мартин-Лёв тщательно проверили текст книги, и я очень благодарен им за их работу. Они отметили некоторое число ошибок и неясностей, а также внесли несколько существенных результатов в теорию.

Я также рад возможности выразить благодарность Скандинавскому страховому акционерному обществу и Шведскому совету исследований в области естественных наук за их финансовую поддержку.

У. Гренандер

Глава 1

ИСТОРИЧЕСКИЕ ПРЕДПОСЫЛКИ И ПРАКТИЧЕСКАЯ МОТИВИРОВКА ВОПРОСА

1.1. Зачем нужно изучать вероятности на общих структурах?

Областью классической теории вероятностей является действительная прямая R^1. Все хорошо известные результаты, вроде центральной предельной теоремы, закона повторного логарифма или закона больших чисел, относятся к случайным величинам, принимающим действительные значения (или векторные значения в R^k). Структурное богатство действительной прямой лежит в основе сложной, но изящной логической конструкции из вероятностных понятий и соотношений.

В математике вообще есть тенденция к обобщению, абстракции, аксиоматике. Очевидно, что это относится также и к теории вероятностей. Со времени опубликования Колмогоровым его эпохальных «Основных понятий теории вероятностей» мы можем использовать вероятностные рассуждения в совершенно общих пространствах без малейшей потери строгости. Правда, при такой общности не всегда можно ожидать результатов реальной математической ценности, но такой общий подход необходим и при более конкретной работе, когда вероятностное пространство наделяется определенными структурными свойствами. Классические результаты показывают, что можно добиться успеха, определяя *алгебраические отношения* в пространстве и связывая их с распределениями вероятностей. Это автоматически заставляет нас обратиться к таким понятиям, как группы, топологические векторные пространства и алгебры. Трудно возражать против такой постановки вопроса: возможна ли на этих более общих алгебраических структурах теория, аналогичная классической теории вероятностей? В следующих главах мы увидим, что иногда это распространение достигается непосредственным и тривиальным обобщением, иногда требуются серьезные усилия для достижения более глубоких результатов, а иногда мы встречаемся с увлекательными проблемами, решение которых пока или совсем неизвестно или известно лишь частично.

То, что теория вероятностей так сильно выросла за последние годы, объясняется, конечно только отчасти, ее внутренней ценностью и ее непосредственной привлекательностью для математиков, но не меньшую роль сыграла и ее полезность, уже эксплуатируемая или потенциальная. Это относится также и к нашему вопросу. Его оправдание состоит не только в желании распространить теорию до ее естественных границ. Существует также некоторое число, казалось бы далеких, задач из физики, теории связи, статистики и т. д., которые приводят нас к рассмотрению вероятностных отношений на алгебраических структурах, не эквивалентных действительной прямой (или пространству R^k). Поскольку все большее число таких задач и результатов приобретает широкую известность, мы можем ожидать возрастающей исследовательской активности и более быстрых успехов, как практических, так и теоретических, в этом направлении.

Мы не сразу погрузимся *in medias res*[1]). Сначала мы вспомним некоторые основные факты и приемы классической теории (в разд. 1.2), затем опишем некоторые задачи, поставленные приложениями (в разд. 1.3), и, наконец, в разд. 1.4 дадим очерк исторического развития теории до ее современного состояния.

1.2. Классические методы и результаты

1.2.1. С разрешения читателя мы начнем с краткого описания некоторых фактов элементарной теории вероятностей.

С действительной прямой R^1 связано некоторое число основных определений теории вероятностей, которые мы будем предполагать известными, так же как и соотношения между ними:

вероятностная мера;
борелевская мера;
случайная величина;
независимость;

---

[1]) In medias res — в суть дела (*лат*).— *Прим. перев.*

различные типы сходимости случайных величин;

сходимость (слабая) распределений вероятностей.

Сейчас для нас наиболее интересны те соотношения, которые используют аддитивные свойства действительных чисел. Пусть P_1, P_2 — две вероятностные меры, скажем, определенные своими функциями распределения:

$$F_i(y)=P_i\{x \mid x\leqslant y\}; \quad i=1,\ 2.$$

Если каждой P_i соответствует случайная величина x_i, $i=1,\ 2$, то сумма $x=x_1+x_2$ имеет распределение вероятностей P, задаваемое функцией распределения $F(x)$. Если x_1 и x_2 независимы, то F есть *композиция* F_1 и F_2, т. е.

$$F(x)=\int_{-\infty}^{\infty} F_1(x-y)\,F_2(dy),$$

или, короче, $F=F_1*F_2$. Существует хорошо известная модификация этой формулы в предположении, что эти распределения абсолютно непрерывны относительно меры Лебега или относительно «считающей меры»[1]) на множестве целых чисел; скажем, если плотности определяются формулами

$$f(x)=\frac{F(dx)}{m(dx)} \quad \text{и} \quad f_i(x)=\frac{F_i(dx)}{m(dx)},$$

то

$$f(x)=\int f_1(x-y)\,f_2(y)\,m(dy).$$

Заметим, что в обоих случаях m есть мера, инвариантная относительно сдвига. Операция композиции коммутативна: $F_1*F_2=F_2*F_1$, и ассоциативна: $(F_1*F_2)*F_3=F_1*(F_2*F_3)$. По индукции определяется $F_1*F_2*\ldots*F_n(x)$. Изучение таких композиций, особенно при больших n, является одной из главных задач теории вероятностей.

[1]) «Считающая (counting) мера» — мера, сосредоточенная на множестве целых чисел и для каждого из них равная единице.— *Прим. ред.*

Отметим ряд элементарных соотношений. Если существуют средние значения

$$m_i = \int_{-\infty}^{\infty} xF_i(dx) = \mathbf{E}x_i; \quad i = 1, 2, \ldots, n,$$

то операция образования среднего значения

$$m = \mathbf{E}x = m_1 + m_2 + \ldots + m_n$$

аддитивна. Если вторые моменты существуют, то дисперсии

$$\mathbf{D}(x_i) = \mathbf{E}(x_i - m_i)^2 = \sigma_i^2$$

удовлетворяют соотношению

$$\mathbf{D}(x) = \sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2$$

при условии, что случайные величины x_i независимы. При этой гипотезе операция образования дисперсии аддитивна, и в силу неотрицательности дисперсия не убывает при добавлении к случайной величине независимого от нее случайного слагаемого. В этом смысле (и во многих других) при композиции распределения «расплываются». Композиция также сглаживает распределения, например в том смысле, что если одно из x_i имеет непрерывное распределение, то и распределение суммы $x = x_1 + x_2 + \ldots + x_n$ также непрерывно.

Наиболее важным аналитическим инструментом для нас является *преобразование Фурье*, или *характеристическая функция*: $\hat{P}(z) = \varphi(z) = \int_{-\infty}^{\infty} e^{ixz} F(dx) = \mathbf{E} \exp(ixz)$, ($z$ действительно). Важность характеристической функции основана на трех ее свойствах:

а) характеристическая функция однозначно определяет вероятностную меру;

б) композиции соответствует обычное умножение: если $P = P_1 * P_2$, то $\hat{P} = \hat{P}_1 \cdot \hat{P}_2$;

в) слабой сходимости последовательности распределений вероятностей к предельному распределению соответствует сходимость характеристических функций к непрерывной предельной функции.

Моменты вероятностной меры P определяются соотношением

$$\alpha_k = \mathbf{E}x^k,$$

если соответствующие интегралы существуют. Они могут быть выражены через производные характеристической функции

$$\alpha_k = \frac{1}{i^k} \varphi^{(k)}(0).$$

Родственными понятиями являются *семиинварианты* (или *кумулянты*)

$$\gamma_k = \frac{1}{i^k} \left[\frac{d^k}{dz^k} \log \varphi(z) \right]_{z=0},$$

которые являются, конечно, линейными комбинациями моментов того же и более низкого порядка. Кумулянты для независимых распределений обладают свойством аддитивности. Некоторые считают, что с моментами и кумулянтами труднее работать и что они определяются менее общим образом, чем характеристическая функция. Возможно, в исследованиях общего характера это и так, но во многих случаях, особенно когда речь идет о предельных теоремах, они очень полезны.

1.2.2. Теперь перейдем к некоторым предельным теоремам. Одной из наиболее старых является *теорема Бернулли*, восходящая к первому этапу развития теории вероятностей: если случайная величина ν имеет биномиальное распределение $B(n, p)$, то частота

$$p^* = \frac{\nu}{n}$$

сходится по вероятности к постоянной p. Или, если записать ν как сумму $\nu = x_1 + x_2 + \ldots + x_n$ независимых случайных величин x_i («индикаторов»), принимающих значения 1 и 0 с вероятностями p и $q = 1 - p$ соответственно, то

$$p^* = \frac{1}{n}(x_1 + x_2 + \ldots + x_n) \to p = \mathbf{E}x_i.$$

Сейчас мы знаем, конечно, что эта формула имеет место при значительно более общих предположениях. Соответствующее утверждение может быть сформулировано различными способами, но одна из возможно наиболее привлекательных формулировок этого *закона больших чисел* принадлежит Хинчину: если случайные величины x_i независимы и одинаково распределены с конечным средним значением m, то

$$\bar{x}=\frac{1}{n}(x_1+x_2+\ldots+x_n)\to m.$$

Эти предельные теоремы могут быть сильно уточнены. Известно (*теорема Муавра — Лапласа*), что распределение нормированной случайной величины

$$\frac{\nu-np}{\sqrt{npq}}$$

сходится к *нормальному* (*гауссовскому*) распределению N (0,1). Или, в более общей формулировке, если величины x_i независимы и одинаково распределены со средним значением m и конечной дисперсией $\sigma^2>0$, то величины

$$\frac{x_1+x_2+\ldots+x_n-nm}{\sigma\sqrt{n}}$$

снова имеют в пределе распределение N (0, 1). Это — простейшая форма *центральной предельной теоремы*. А какое предельное распределение будет иметь соответствующим образом нормированная сумма

$$x=\frac{x_1+x_2+\ldots+x_n-a_n}{b_n}$$

в еще более общей ситуации, когда x_i независимы и одинаково распределены, а относительно моментов не делается никаких предположений? Известно (Хинчин), что в этом случае возможные предельные распределения *устойчивы*. Соответствующие функции распределения F таковы, что для любых $a_1>0$, $a_2>0$ и произвольных b_1 и b_2 существуют постоянные $a>0$ и b, для которых

$$F(a_1x+b_1)*F(a_2x+b_2)=F(ax+b).$$

Устойчивые распределения полностью описаны в терминах характеристических функций (см. замечания 1.2.2.$_1$).

Существуют другие предельные теоремы, не укладывающиеся непосредственно в эту схему. Укажем на известный *пуассоновский случай*, в котором биномиальное распределение стремится к распределению Пуассона, когда $n \to \infty$, $np \to \lambda$. В этом случае распределение x_i не остается фиксированным при $n \to \infty$, но изменяется, так как параметр p становится все меньше и меньше. Чтобы сформулировать это в более общих терминах, нужно рассмотреть *треугольную таблицу* случайных величин

$$\begin{array}{l} x_{11} \\ x_{21},\ x_{22} \\ x_{31},\ x_{32},\ x_{33} \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \end{array}$$

Случайные величины в каждой строке предполагаются независимыми и одинаково распределенными. Второе условие может быть заменено условием, что ни одно слагаемое в сумме $x_{n1} + x_{n2} + \ldots + x_{nn}$ не дает вклада, соизмеримого по величине со всей суммой. Иначе мы могли бы, конечно, получить любое предельное распределение. Таким условием может, например, являться условие

$$\lim_{n\to\infty} \sup_{1\leqslant k\leqslant n} \mathbf{P}\{|x_{nk}| \geqslant \varepsilon\} = 0$$

для любого $\varepsilon > 0$. В этом случае класс возможных предельных распределений совпадает с классом *безгранично делимых* распределений. Распределение вероятностей P называется безгранично делимым, если для любого натурального числа n существует такое распределение P_n, что

$$P = P_n^{n*} = \underbrace{P_n * P_n * \ \ldots\ * P_n}_{n \text{ раз}}.$$

Распределение вероятностей безгранично делимо тогда и только тогда, когда его характеристическая функция имеет вид

$$\varphi(z) = \exp\left\{iaz + \int \left[e^{iuz} - 1 - \frac{iuz}{1+u^2}\right] \frac{1+u^2}{u^2}\, G(du)\right\}$$

(*представление Леви — Хинчина*), где a — действительная постоянная, а G — неубывающая функция ограниченной вариации.

Известны критерии, позволяющие непосредственно судить об асимптотическом поведении суммы случайных величин и выраженные в терминах простых свойств индивидуальных функций распределения. Это составляет в известном смысле законченную теорию, изложенную с замечательной ясностью и точностью в книге Гнеденко и Колмогорова «Предельные распределения для сумм независимых случайных величин», представляющей собой *monumentum aere perennius* [1]) в литературе по теории вероятностей.

1.2.3. Оказывается весьма полезным построение случайного процесса, соответствующего данной предельной теореме. Эта идея, выдвинутая и систематически проведенная в книге Хинчина «Асимптотические законы теории вероятностей», основана на понятии *случайного процесса с независимыми приращениями*. Так называются процессы $x(t)$, для которых приращения $\xi_1 = x(t_2) - x(t_1)$, $\xi_2 = x(t_3) - x(t_2)$, ..., $\xi_n = x(t) - x(t_n)$ являются независимыми случайными величинами при $0 = t_1 < t_2 < \ldots < t_n < t$. Особенно важны *однородные* процессы, для которых распределение P_h любого приращения $x(t+h) - x(t)$, $h \geqslant 0$, зависит только от h (но не от t). Такие процессы называются *непрерывными по вероятности*, если $P_h \to \delta_0$ (слабо) при $h \downarrow 0$; символом δ_x мы всегда обозначаем вырожденное распределение, при котором вся масса сосредоточена в точке x.

Полагая $x(0) = 0$, можно записать $x(t)$ как сумму независимых случайных величин

$$x(t) = \xi_1 + \xi_2 + \ldots + \xi_n.$$

Если $x(t)$ обладает свойствами, определенными выше, то очевидно, что $x(t)$ должно обладать безгранично делимым распределением при любом $t > 0$. Легко также показать, что характеристическая функция $\varphi(z, t)$ величины $x(t)$ должна иметь вид

$$\varphi(z, t) = [\psi(z)]^t,$$

где $\psi(z)$ есть характеристическая функция (скажем, в представлении Леви — Хинчина) некоторого безгранично делимого распределения, а именно распределения $x(1)$.

[1]) Неразрушимый, вечный памятник (Гораций, Оды, кн. III, 30).— *Прим. перев.*

Теперь соответствие между процессом и связанной с ним предельной теоремой интуитивно ясно, хотя его и не так легко установить с полной строгостью и во всех подробностях. Локальные свойства выборочных функций процесса (непрерывность, возможная величина скачков и т. д.) определяются функцией G в представлении Леви — Хинчина. Укажем только, что если выборочные функции непрерывны почти наверное (в предположении сепарабельности процесса), то G постоянна, за исключением $u=0$, так что характеристическая функция имеет вид

$$\varphi(z,\ t)=\exp\left[t\left(imz-\frac{\sigma^2}{2}z^2\right)\right],$$

соответствующий нормальному процессу. Подробное изложение читатель найдет в книге Дуба [1]. Пуассоновский процесс является другим легко разбираемым примером.

С другой стороны, известно, что функция G предельного закона может быть выражена через функции распределения слагаемых x_{nk}. Это устанавливает логическую связь между поведением этих слагаемых в предельной теореме и локальными свойствами выборочных функций соответствующего случайного процесса с независимыми приращениями.

В математическом аппарате, используемом для доказательства классических предельных теорем, анализ Фурье занимает, конечно, центральное место, но в некоторых специальных случаях проще использовать преобразование Лапласа, моменты и т. п. Иной подход используется при установлении соответствия, о котором говорится в данном разделе. Выводится функциональное уравнение, описывающее поведение распределений соответствующего процесса. Это может быть уравнение теплопроводности, уравнение Фоккера — Планка или уравнение более общего вида. Затем рассматривают распределения накопленных сумм, о которых идет речь, и стараются показать, что они удовлетворяют функциональному уравнению, в некотором смысле близкому к первому уравнению, и, наконец, применяют рассуждение по непрерывности (см. замечания $1.2.2._2$).

В следующих главах уделяется большое внимание развитию аппарата для изучения предельных теорем для независимых величин и процессов с независимыми приращениями. Мы будем, кроме того, все время руководствоваться тем, что мы знаем из классической теории.

1.3. Практические предпосылки теории

Чувствуя эстетическую привлекательность общей и законченной математической теории, автор все же считает, что сильнейшим стимулом к продолжению работы в вероятностной теории алгебраических структур будут являться приложения. Уже сейчас разнообразие и число практических вопросов, приводящих к задачам, которые мы имеем в виду, указывают на необходимость такой теории. Здесь мы только укажем несколько типичных случаев.

1.3.1. Пусть в нашем распоряжении есть некоторый метод получения случайных чисел. В десятичной записи они имеют вид

$$x = 0,\ x_1 x_2\ \ldots\ x_k.$$

Каждому из 10^k различных возможных чисел соответствует некоторая вероятность, определяемая распределением вероятностей, которое может быть известно нам только частично. Пусть получена последовательность независимых чисел $x^{(1)}, x^{(2)}, \ldots, x^{(n)}$. Скомбинируем их следующим образом. Сложим эти числа и выделим дробную часть ξ_n суммы:

$$x^{(1)} + x^{(2)} + \cdots + x^{(n)} = \text{Натуральное число} + \xi_n.$$

Если n велико, то можно надеяться, что распределение ξ_n стабилизируется около равномерного распределения (по возможным значениям), т. е. около полностью известного распределения. Очевидно, что здесь мы имеем дело с *циклической группой* Z порядка 10^k. Эта задача настолько проста, что она может быть решена непосредственно совершенно элементарными методами. Однако в этом решении скрыт зародыш более общей идеи, которая будет обсуждаться в дальнейшем.

1.3.2. Пусть ξ_t — стационарная нормальная последовательность со средним 0 и корреляционной функцией $r_h = \mathbf{E}\,\xi_t \xi_{t+h}$; $t, h = 0, \pm 1, \pm 2, \ldots$. Если ξ_t проходит через фильтр, мгновенно реагирующий на входной сигнал, так что на выходе мы получаем достаточно хорошую функцию $g(\xi_t)$ от ξ_t, то мы можем изучать свойства функции на выходе. Как это часто бывает в случае стационарных процессов, здесь предпочтительнее работать со спектрами, а не со вторыми моментами.

Вычисления приводят нас к представлению спектральной функции выхода в виде линейной комбинации спектральных функций, соответствующих корреляционным функциям r_h^ν; $\nu = 0, 1, 2, \ldots$ На языке частот λ это равносильно образованию последовательных композиций исходной спектральной функции $F(\lambda)$:

$$\begin{array}{l} r_h \longleftrightarrow F(\lambda) \\ r_h^2 \longleftrightarrow F * F(\lambda) \\ r_h^3 \longleftrightarrow F * F * F(\lambda) \\ \ldots\ldots\ldots\ldots \end{array}$$

Хотя некоторые из этих композиций и могут быть вычислены непосредственно, ясно, что для композиций более высокого порядка можно использовать аппроксимации. Можно записать $F = \sigma^2 G$, где G есть обычная (нормированная) функция распределения, а σ^2 — дисперсия на входе. Тогда имеем

$$F^{n*} = \sigma^{2n} G^{n*}.$$

В интегралах, определяющих F^{n*} и G^{n*}, мы отождествляем частоты, отличающиеся множителем 2π. Другими словами, мы имеем дело со сложением в группе T^1 на одномерном торе, G определяет распределение вероятностей на этой группе, а G^{n*} есть результат «сложения» n независимых и одинаково распределенных элементов группы. Мы получим полезную аппроксимацию, если покажем, что G^{n*} сходится к некоторому предельному распределению. Снова почти очевидно, что такое предельное распределение, если оно существует, должно обладать свойством равномерности (т. е. инвариантности, см. замечания 1.3.2).

Сказанное выше может быть почти слово в слово повторено применительно к группам T^k на k-мерном торе; в принципе решение столь же просто и при бóльших k.

Отметим, что до сих пор нам встречались только коммутативные группы.

1.3.3. Рассмотрим теперь физическую систему (жидкость, газ) с частицами, совершающими броуновское движение. Для конкретности предположим, что среднее значение сноса равно нулю и что движение изотропно. В эту систему погружен шар с фиксированным центром; движение шара связано силами трения с движением окружающей среды (см. замечания 1.3.3).

В заданном интервале времени (t_1, t_2) шар вращается, и это вращение мы обозначим O_1. Полезно представить себе одну фиксированную систему координат и одну подвижную, перемещающуюся вместе с шаром. В координатной форме O_1 выражается ортогональной матрицей. Так или иначе, O_1 есть случайное вращение. В следующем интервале времени (t_2, t_3) происходит новое вращение O_2, стохастически независимое от первого при некоторых условиях (отсутствие инерции...), и т. д. Наконец, в интервале (t_n, t_{n+1}) происходит вращение O_n. Полное вращение за время (t_1, t_{n+1}) будет тогда $R_n = O_1 O_2 \ldots O_n$. Чтобы изучить поведение R_n при больших значениях n, мы должны рассмотреть предельные вероятностные законы на *ортогональной группе*. Здесь мы сталкиваемся с интересным обстоятельством. В предыдущих примерах группы были коммутативными, но при композиции вращений порядок выполнения операций существен. Это свойство *некоммутативности* вызывает соответствующее усложнение используемого математического аппарата.

В этом примере броуновское движение действовало на очень простое физическое тело, а именно на шар. Нетрудно представить реальные ситуации, в которых шар заменяется другим твердым или упругим телом или в которых рассматриваются не механические, а электрические системы. Каждый раз, когда мы можем выписать дифференциальные уравнения, управляющие поведением системы, мы будем приходить к аналогичным, хотя, может быть, и более сложным, задачам. Объектом изучения будут *группы Ли*, целью — введение вероятностных структур на таких группах. В случаях когда эти группы *компактны*, можно надеяться отыскать предельные (равномерные) распределения, и читателю может прийти мысль о *мере Хаара*. И он будет прав; это понятие будет играть здесь основную роль. Конечно, при этом несущественно, что группы являются группами Ли, важно, что они компактны.

1.3.4. *Неупорядоченная линейная структура* есть одномерная цепь частиц (слово это может пониматься очень широко), в которой смежные частицы связаны упругими силами и в которой частицы или силы имеют случайные характеристики. Часто можно представлять, что действие частицы на следующую в цепи описывается *матрицей передачи*. В простейшем случае это квадратные матрицы второго порядка. Они комбинируются с помощью обычного умножения матриц. Если эти *стохастические матрицы* независимы, то положение аналогично описанному выше. Однако вычисления показывают, что соответствующие группы не всегда являются компактными. Это приводит к крайнему возрастанию степени математических трудностей при попытке найти математический инструмент для решения этой стохастической задачи. Как мы увидим в дальнейшем, существуют новые и фундаментальные проблемы уже в *формулировании* соответствующих предельных теорем (см. замечания 1.3.4).

Представим себе цепь приборов совершенно общего физического характера. Если действие такого прибора может быть описано матрицей передачи, то мы все еще в знакомой обстановке. Пусть цепь описывается линейным (разностным) уравнением

$$u_{n+p}+a_n^{(p-1)}u_{n+p-1}+\cdots+a_n^{(0)}u_n=v_n,$$

в котором правая часть может быть детерминированной или случайной, возможно, тождественно равной нулю, а коэффициенты $a_n^{(j)}$ случайны. Рассмотрим векторы

$$\xi_n=\begin{Bmatrix} u_{n+p} \\ u_{n+p-1} \\ \vdots \\ u_{n+1} \\ u_n \end{Bmatrix}.$$

Тогда $\xi_{n+1}=M_n\xi_n+\eta_n$, где M_n есть стохастическая матрица. В частности, если все η равны нулю, мы получаем модель, тесно связанную с предыдущей.

Так как стохастические свойства матрицы передачи можно выбрать различными способами, то можно ожидать большого числа различных, хотя и родственных друг другу, результатов. В качестве примера упомянем случай волноводов со стохастической неоднородностью. Используя аппроксимацию дискретной цепью, можно описать этот случай в терминах (комплексных) коэффициентов отражения. Оказывается, что этот путь приводит к изучению распределений вероятностей на открытом единичном круге $|z| < 1$, причем закон композиции определяется соотношением

$$z_1 \circ z_2 = \frac{z_1 + z_2}{1 + \bar{z}_1 z_2}.$$

Множество дробно-линейных преобразований

$$\frac{z_1 + z}{1 + \bar{z}_1 z}$$

образует хорошо известную группу преобразований, оставляющих инвариантной единичную окружность. Наша задача, таким образом, аналогична предыдущей и заключается в изучении распределения вероятностей композиции $z_1 \circ z_2 \circ \ldots \circ z_n$ большого числа независимых и одинаково распределенных групповых элементов z_ν.

Если разностное уравнение нелинейно, то матрицы нужно заменить нелинейными операторами. Мы снова можем прийти к стохастическим группам, во всяком случае, если эти операторы несингулярны. Однако, может быть, их стоит рассматривать как элементы *стохастической алгебры*, и этот подход имеет очень интересные приложения, как мы увидим в последней главе. Об этом рано говорить, но достаточно указать на увлекательную проблему определения спектральных свойств элементов таких стохастических алгебр.

Может случиться, что пространственно-временна́я шкала физической задачи такова, что дифференциальное уравнение дает лучшее описание физической картины, чем приведенное выше разностное уравнение. Ему будут соответствовать случайные процессы со значениями из группы, алгебры и т. п.; по всей вероятности, это будут процессы с «независимыми приращениями». Особенно важны бесконечно малые преобразования, так как они допускают простую и конкретную физическую интерпретацию.

Заметим, что до сих пор все пространства, на которых рассматривались распределения вероятностей, были *локально компактными*.

1.3.5. Обратимся теперь к задаче обнаружения сигнала. Мы наблюдаем случайный процесс $x(t)$, $t \in (0, T)$. Для простоты предположим, что $x(t)$ есть нормальный процесс с известной корреляционной функцией, но с двумя гипотетическими средними H_0: $\mathbf{E}\, x(t) \equiv 0$; H_1: $\mathbf{E}\, x(t) = m(t)$; $t \in (0, T)$. Нулевая гипотеза соответствует случаю, когда сигнала нет, а есть только шум. При альтернативной гипотезе $x(t)$ содержит сигнал. Мы хотим найти критерий для различения этих двух гипотез.

Подробные статистические задачи решаются сейчас следующим образом. Двум гипотезам соответствуют две вероятностные меры P_0 и P_1 в некотором вероятностном пространстве Ω. Вычисляется производная Радона — Никодима $p(\omega) = P_1(d\omega)/P_0(d\omega)$, если P_1 абсолютно непрерывна относительно P_0, и образуется критическая область Неймана — Пирсона

$$W = \{\omega \mid p(\omega) \geqslant c\},$$

где константа c выбирается так, чтобы вероятность $P_0(W)$ имела приемлемое значение. В этом случае мы имеем дело с линейной задачей, и можно найти $p(\omega)$ простым применением методов гильбертова пространства. При таком подходе мы рассматриваем процесс как континуум случайных величин, каждая из которых представляется точкой в гильбертовом пространстве. Следовательно, мы не специфицируем индивидуальные выборочные функции. С некоторой точки зрения более привлекательным (хотя, возможно, не более практичным) было бы начать с индивидуальных выборочных функций, оперировать с ними по соответствующей схеме и получить критерий для обнаружения сигнала.

Это можно сделать, предполагая, что совокупность выборочных функций образует некоторое функциональное пространство. Можно использовать некоторое *банахово пространство.* При этом нужно ввести вероятностные меры на таких пространствах и изучить свойства этих мер. Заметим, что эти пространства, вообще говоря, не являются локально компактными (что приводит к некоторым усложнениям), но зато они имеют линейную структуру (что, конечно, упрощает дело).

Чтобы более ясно показать, чего можно добиться с помощью этого подхода, предположим, что делаются повторные выборки из того же процесса, что и раньше. Обозначим наблюденные выборочные функции $x_1(t)$, $x_2(t)$, ..., $x_n(t)$ и образуем среднее:

$$\bar{x}_n(t) = \frac{1}{n}[x_1(t) + x_2(t) + \cdots + x_n(t)].$$

Это среднее также принадлежит линейному пространству, и нас может интересовать вопрос о его сходимости (в топологии того же пространства) при $n \to \infty$. Было бы, конечно, желательно, чтобы $\bar{x}_n(t)$ сходилось к $m(t) = \mathbf{E}\, x(t)$ в смысле (сильной) сходимости по вероятности:

$$\lim_{n\to\infty} \mathbf{P}\{\| m(t) - \bar{x}_n(t) \| > \varepsilon\} = 0 \text{ при любом } \varepsilon > 0,$$

или в смысле (сильной) сходимости почти наверное:

$$\mathbf{P}\{\lim_{n\to\infty} \| m(t) - \bar{x}_n(t) \| = 0\} = 1,$$

или в смысле какой-нибудь другой сходимости. Подобные формы *закона больших чисел в линейных пространствах* были бы очень полезны.

Точно так же было бы желательно получить различные формы центральной предельной теоремы и т. п. в стохастических линейных пространствах. Достаточно будет одного примера. Предположим, что мы делаем выборку из непрерывной одномерной популяции с известной функцией распределения. Известно, что можно преобразовать это распределение в равномерное распределение на единичном интервале (0, 1). Образуем эмпирическую функцию распределения

$$F_n(x) = \frac{1}{n} [\text{число наблюдений } \leqslant x].$$

Тогда функцию

$$y_n(x) = \sqrt{n}\, [F_n(x) - x]$$

можно рассматривать как элемент гильбертова пространства $L_2(0, 1)$ всех функций с интегрируемым квадратом на (0, 1). Это определяет распределение вероятностей P_n на пространстве $L_2(0, 1)$. Хотелось бы доказать, что последовательность $\{P_n\}$ сходится (в каком смысле?) при $n \to \infty$ к некоторому предельному распределению P. Было бы очень интересно вывести, что

$$\mathbf{P}_n\{| y_n(x) | \leqslant c\} \to \mathbf{P}\{| y(x) | \leqslant c\};$$

это связано, конечно, с критериями Колмогорова и Смирнова (см. также замечания $1.2.2._2$).

1.3.6. В конце разд. 1.3.4 мы говорили о стохастических алгебрах. Тогда мы имели в виду только локально компактные алгебры, но теперь естественно обратиться к *банаховым алгебрам*, комбинируя линейную структуру вероятностного пространства с линейной структурой алгебры. Это естественно делать в тех приложениях, где ограниченный стохастический оператор действует линейно на банаховом пространстве. Неограниченные операторы, которые появляются, например, в связи со стохастическими дифференциальными операторами, находятся за пределами нашего изложения.

1.4. Исторические предпосылки

1.4.1. Всегда бывает трудно проследить развитие математических идей в обратном порядке вплоть до момента их возникновения; но с некоторым усилием и при желании можно, несмотря на неясность и неопределенность задачи, добиться успеха, продвинувшись при этом далеко назад во времени. Наш случай не является исключением из этого правила. Хотя большинство результатов современной теории относится к пятидесятым годам и особенно продуктивными были именно последние годы, некоторые идеи возникли значительно раньше.

В классической теории вероятностей, в таких ее разделах, как предельные теоремы и т. п., рассматривались распределения на действительной прямой или на решетках, расположенных на ней, но уже на ранних этапах развития теории выяснилось, что она могла бы быть распространена на распределение на плоскости R^2 или в евклидовом многомерном пространстве R^k. Возьмем, например, закон больших чисел. Стоило только установить его для случайных величин, принимающих числовые действительные значения, как сразу же оказалось возможным сформулировать и доказать его, практически тем же способом, для величин, принимающих векторные значения. При этом математически вносится очень мало нового. Другие предельные теоремы, возможно, несколько труднее обобщить на R^k, но в общем положение остается прежним: это не очень трудно, может быть, не очень интересно, но подчас весьма полезно.

1.4.2. Переходя к бесконечномерным векторным пространствам, мы сталкиваемся с совершенно другой ситуацией. Здесь требуются более серьезные усилия для перенесения классических результатов или для отыскания новых результатов, занимающих место старых, уже известных. В серии работ Фреше ([1] и др.) обоснована необходимость изучения *теории вероятностей в топологических пространствах* различной степени общности. Его предложение — весьма естественное для одного из создателей современного функционального анализа — не встретило сначала широкого отклика, и прошло некоторое время, прежде чем специалисты по теории вероятностей осознали возможности этого направления. Исключением в этом медленном развитии является одна ранняя работа Колмогорова [2]. Уже в 1935 г. он предложил изучать теорию вероятностей на банаховых пространствах с помощью того, что он назвал *характеристическим функционалом*. Пусть X есть банахово пространство с элементами x; X^* — двойственное пространство линейных (ограниченных) функционалов $x^* = x^*(x)$ и P — распределение вероятностей на X. На время отложим вопрос о том, как в этом пространстве следует ввести распределение P. Колмогоров определяет характеристический функционал как

$$\hat{P}(x^*) = \mathbf{E} \exp[ix^*(x)];$$

это комплекснозначная функция на X^*. Она обладает по крайней мере некоторыми из главных свойств характеристической функции. Характеристический функционал однозначно определяет распределение вероятностей (теорема единственности), композиция соответствует обычному умножению комплекснозначных функций. Характеристический функционал является положительно определенной функцией. Однако другие вопросы, вроде аналога теоремы непрерывности, лежат значительно глубже и потому в течение некоторого времени оставались открытыми.

Понятие характеристического функционала предлагалось рядом авторов в связи с различными задачами: Бох нером [1], Ле Камом [1] и другими. Оно является неизбежным инструментом при изучении стохастических линейных пространств.

1.4.3. Систематическое изучение всех этих проблем началось около 1953 г. В фундаментальной работе Мурье был более подробно изучен анализ Фурье и был сделан большой шаг в теории предельных теорем. Были получены некоторые варианты закона больших чисел и результат, соответствующий центральной предельной теореме в гильбертовом пространстве. Последний был завершен в работе Мурье и Форте [2]. Нормальное распределение определялось другими, но эквивалентными способами, например как распределение, для которого все линейные функционалы имеют нормальное (действительное или комплексное) распределение.

Мурье рассматривала также построения теории меры, которые должны быть использованы в банаховом пространстве. Распределения вероятностей вводятся с помощью всех (скалярных) случайных величин $x^*(x)$, где x^* пробегает двойственное пространство X^*. Эти функции предполагаются измеримыми и тогда их можно считать обычными случайными величинами. Несколько отличный подход был использован Ганшем [1], который предположил вместо этого, что P есть борелевская мера. В то время как Мурье определяла распределения вероятностей на полупространствах $\{x \mid x^*(x) \leqslant c\} \subset X$, Ганш положил в основу открытые множества. Для сепарабельного пространства эти два метода приводят к одному и тому же результату, и только в несепарабельном случае разница существенна.

Поскольку линейные функционалы измеримы, можно определить *математическое ожидание* случайного элемента в банаховом пространстве следующим образом (см. замечания 1.4.3). Если существует такой элемент $m \in X$, что равенство

$$x^*(m) = \int\limits_X x^*(x)\, P(dx)$$

имеет смысл и выполнено для всех $x^* \in X^*$, то m называется математическим ожиданием P и обозначается $m = \mathbf{E}x$. Это не что иное, как интеграл Петтиса, примененный к нашей ситуации. В банаховом пространстве сохраняются многие свойства обычного математического ожидания. Это определение необходимо для закона больших чисел.

Моменты второго порядка (или лучше операторы ковариации) могут быть определены в стохастическом гильбертовом пространстве. Моменты более высокого порядка, по-видимому, до сих пор не использовались.

Теперь заложены основы теории *стохастических банаховых пространств*, но имеют место по крайней мере два неприятных пробела. Одним из них является уже отмеченная трудность, возникающая при попытке распространения теоремы непрерывности. Пусть мы имеем последовательность вероятностных мер P_1, P_2, ... на X. В обычной терминологии P_n *слабо сходится* к предельному распределению P, если

$$\lim_{n\to\infty}\int_X f(x)\,P_n(dx)=\int_X f(x)\,P(dx)$$

для любой непрерывной и ограниченной комплекснозначной функции $f(x)$. Если это так, то, конечно, $\hat{P}_n(x^*)\to\hat{P}(x^*)$, $x^*\in X^*$, где значок «$\wedge$» обозначает преобразование Фурье (характеристический функционал). Это сразу видно, если положить $f(x)=\exp[ix^*(x)]$. Обратно, пусть известно, что $\hat{P}_n(x^*)$ сходится. Можно ли тогда утверждать, что P_n слабо сходится к некоторому распределению вероятностей? Простые примеры показывают, что, вообще говоря, это не так, и этот неприятный факт объясняется тем, что распределенная масса может «растекаться» во все более высокие размерности (представьте себе гильбертово пространство со счетным множеством координатных осей). Чтобы справиться с этим, нужно наложить какое-то условие компактности на P_n. Это может быть сделано различными способами, в частности для гильбертова пространства требуемые условия установлены Прохоровым [1].

Другой пробел заключается в следующем. Предположим, что $\varphi(x^*)$ есть комплексная функция на X^*, непрерывная по x^*, удовлетворяющая условиям $\varphi(0)=1$, $\varphi(-x^*)=\overline{\varphi(x^*)}$, и положительно определенная. Существует ли такое распределение вероятностей P на X, что $\varphi(x^*)=\hat{P}(x^*)$? И снова ответ — не всегда. Можно заподозрить, что все дело в понятии «непрерывности». Мы интерпретировали ее как «сильную непрерывность». Если вместо этого рассматривать слабую сходимость, то мы все еще не придем к правильному утверждению. Колмогоровым [3] для гильбертова пространства было показано, что для этого должна быть использована другая сходимость (S-сходимость), и это дает частичное решение задачи. Другое, более общее, но несколько более сложно формулируемое условие было дано Мурье [1].

Эта теория быстро развивается.

1.4.4. Можно разумно наделить вероятностной структурой и другие пространства. Если пространство обладает некоторыми линейными свойствами, то положение аналогично случаю банахова пространства. Гельфанд [1] и К. Ито [2] ввели вероятностные меры во множестве обобщенных функций *распределений Шварца*. Эта работа была продолжена Ульрихом [1] и другими. Анализ Фурье по-прежнему остается, конечно, главным инструментом.

Другой линией развития являются общие метрические пространства. Несмотря на то что в таких пространствах можно определить распределения вероятностей и даже средние значения и т. п., ясно, что мы не можем получить результаты, подобные описанным выше, не определив сложения или какой-нибудь другой бинарной операции. Поэтому это направление выходит за пределы настоящей книги и не будет обсуждаться в дальнейшем.

1.4.5. Если какое-нибудь понятие должно быть выделено в нашей теме как центральное, то это, безусловно, понятие *стохастической группы*. Линейные пространства являются группами относительно сложения, но, кроме этого, обладают рядом других свойств. Можно получить другие ценные результаты, если вместо линейности предположить (локальную) компактность топологической группы.

В литературе существуют несколько изолированных, но важных ранних исследований в этом направлении. Уже в 1928 г. физические рассмотрения привели Перрэна [1] к рассмотрению стохастической группы на группе вращений в трехмерном пространстве. Несмотря на специальный характер этого исследования, оно содержит весьма общие идеи.

То же можно сказать и о работе Леви [1], в которой подробно рассматривается группа вращений окружности.

Прорыв был сделан в 1940 году работой Кавада и К. Ито [1]. Эта работа, которая до последнего времени была почти незамеченной, значительна и по своим результатам и особенно по своим методам. Авторы исследовали поведение композиций P^{n*} на компактных группах и установили, что возможные предельные распределения представляют собой меры Хаара на некоторых подгруппах. Для того чтобы доказать это, они использовали анализ Фурье, сильно опирающийся на теорию унитарных представлений компактных групп. Рассмотрим полную систему неприводимых унитарных представлений

$$U_0(g) \equiv I, \quad U_1(g), \quad U_2(g), \; \ldots \; .$$

Все эти $U_j(g)$ могут быть заданы как унитарные матрицы в конечномерных пространствах. Преобразование Фурье распределения P определяется тогда как

$$\hat{P}_j = \int_G U_j(g)\, P(dg); \quad j = 0,\ 1,\ 2,\ \ldots\ .$$

Так как эти преобразования Фурье ведут себя подобно характеристическим функциям, то они хорошо приспособлены для доказательства предельных теорем и т. п., что и было использовано К. Ито и Кавада.

Эти результаты были заново открыты, модифицированы и обобщены рядом авторов, среди которых нужно указать, в частности, Стромберга [1], Клосса [1] и Урбаника [1]. Конечные группы изучались Воробьевым [1], Дворецким и Вольфовицем [1] и Бёге [1]; последний изучал безгранично делимые законы.

К настоящему времени теория компактных стохастических групп уже построена; возможно, нельзя утверждать, что она закончена, но она дает ответ на некоторые из основных вопросов. Оглядываясь назад, можно было бы думать, что компактность настолько упрощает задачу, что она может быть решена без помощи тонкого аналитического аппарата. Однако это было бы недооценкой исходных концептуальных трудностей. Полученные результаты указывают также путь для дальнейшего развития.

1.4.6. Если предположить, что группа только локально компактна, то мы встретимся с дополнительными трудностями. Понятие преобразования Фурье обобщается достаточно просто, но его основные свойства не устанавливаются автоматически. Результаты в этом направлении были получены в работах Гренандера [3, 4], там же приведены некоторые результаты, установленные с помощью анализа Фурье.

Для этой цели им была использована теория представлений локально компактных групп, развитая Годеманом, Наймарком и др. Это привело к теоремам, аналогичным закону больших чисел, и центральной предельной теореме. Было также показано, что существуют простые ситуации, в которых результаты не похожи на те, которые имеют место в классической теории. В этой весьма общей ситуации наши знания менее полны, но контуры теории уже различимы.

Оператор

$$Tf(g) = \int\limits_G f(gh)\, P(dh)$$

играет важную роль. Его спектральные свойства были изучены Гренандером [2] и Кестеном [1].

Интересный частный случай представляет собой группа Ли. Хант уже в 1956 г. показал, что однородные процессы на такой группе могут быть полностью охарактеризованы. Хант сделал это, выведя общее выражение для *инфинитезимального порождающего оператора*

$$Af(g) = \lim_{h \downarrow 0} \frac{T_h - I}{h} f(g),$$

где

$$T_t f(g) = \int\limits_G f(gh)\, P_t(dh)$$

и где функция $f(g)$ достаточно гладкая. Аналогичными методами Вен [1] показал возможность вывода нескольких важных предельных теорем.

Возвращаясь к общему случаю, подчеркнем важность понятия *идемпотентной меры*. Вероятностная мера M называется идемпотентной, если $M^{2*} = M * M = M$. Пусть P

есть вероятностная мера на группе, а Q есть предельная мера $P^{n*} \to Q$. Тогда $Q^{2*} = \lim_n P^{n*} * P^{n*} = \lim_n P^{n*} = Q$, и мера Q должна быть идемпотентной. Обратно, любая идемпотентная мера Q является предельной: $Q = Q^{n*} = \lim_{n\to\infty} Q^{n*}$. Класс возможных предельных распределений совпадает с классом идемпотентных распределений, что и определяет их важность для нас. Идемпотентные меры были изучены Венделем [1] и др. при различных предположениях относительно положенной в основу алгебраическо-топологической структуры. Типичным идемпотентом является нормированная мера Хаара на компактной подгруппе.

Если Q есть идемпотент, то ее носитель, т. е. наименьшее множество, содержащее всю распределенную массу (более аккуратное определение будет дано ниже), должен быть подгруппой. Чтобы почувствовать это, представим себе конечное распределение с точечиыми массами p_g, $g \in G$. Имеем

$$p_g = \sum_{ab=g} p_a p_b.$$

Введем множество $\Gamma = \{g \mid p_g > 0\}$. Если p_a, $p_b > 0$, то $p_{ab} > 0$, так что $\Gamma^2 \subset \Gamma$. Но так как Γ конечно, отсюда следует, что Γ образует подгруппу. Мы увидим в основном тексте, что это имеет место и в более общей обстановке.

Тип групп, более близких к линейиым пространствам, образуют группы с *однозначно определенными корнями* n-й степени: каждому натуральному n и каждому элементу $g \in G$ должен соответствовать одиозначно определенный элемент $h \in G$, такой, что $h^n = g$. Такие стохастические группы рассматривались Гренаидером [4]. На таких группах предельные теоремы прииимают особенно приятный вид.

1.4.7. Некоторые авторы изучали *стохастические полугруппы*; их исследования показали, где и насколько существенны групповые свойства для справедливости результатов, о которых шла речь выше. До сих пор подробно исследовались только компактные полугруппы.

Хьюит и Цуккерман [1] изучали конечные коммутативные полугруппы, используя полухарактеры для своего определения преобразования Фурье распределения вероятностей. Полухарактер $x(s)$ есть комплекснозначная функция $x(s) \not\equiv 0$, такая, что $x(st) = x(s)\,x(t)$ для любых элементов s и t полугруппы. Преобразование Фурье определяется тогда как

$$\hat{P}(x) = \sum_s p_s x(s),$$

и оно обладает некоторыми известными свойствами преобразований Фурье. Среди других результатов эти авторы дали критерий сходимости P^{n*}.

Розенблат [1] (см. также Хебл и Розенблат [1]) изучал предельные теоремы на компактных полугруппах и среди прочих результатов показал, что средние

$$\pi_n = \frac{1}{n}(P + P^{2*} + \cdots + P^{n*})$$

всегда сходятся к некоторой предельной мере π при $n \to \infty$. Мера π является идемпотентом, и ее носитель есть минимальный двусторонний идеал полугруппы.

Так же, как появление подгрупп в случае группы, естественно появление ядра (минимального двустороннего идеала) при изучении предельных теорем и идемпотентных мер. Легко дать этому эвристическую мотивировку. Предположим, что P^{n*} сходится к Q. Случайный элемент $\gamma_n = g_1 g_2 \ldots g_n$ принимает значения из носителя $s(P^{n*})$, и мы можем ограничиться замкнутой полугруппой S, натянутой на эти носители, $n = 1, 2, \ldots$. Если I есть двусторонний идеал полугруппы S и n велико, то значения γ_n будут элементами из S, близкими к I. Но если мы попадаем в I, мы там остаемся. Рассмотрим теперь

$$\gamma_{rn} = g_1 \ldots g_n g_{n+1} \ldots g_{2n} \ldots g_{(r-1)n} \ldots g_{rn},$$

разбитое на r больших групп. С подавляющей вероятностью (если r также велико) по крайней мере одна из этих групп представляет элемент, близкий к I. Следовательно, мы можем ожидать, что предельное распределение имеет своим носителем I, по-видимому, мы даже можем надеяться, что $s(Q)$ будет наименьшим I, т. е. ядром K.

1.4.8. Другим направлением исследования являются стохастические алгебры. Там мы имеем две главные операции — сложение и умножение. Гренандер [2, 3] исследовал отношения между аддитивными и мультипликативными процессами и их влияние на соответствующие предельные теоремы. Если, в частности, объектом изучения является банахова алгебра, то мы приходим к естественной связи с тем, что мы знаем относительно теории вероятностей в банаховых пространствах.

Важный частный случай представляют собой квадратные матрицы порядка k. Ферстенберг и Кестен [1] показали, что предел

$$\lim_{n\to\infty} \frac{1}{n} \mathbf{E} \log \| M_1 M_2 \ldots M_n \|$$

существует и что, если он конечен, то равен пределу

$$\lim_{n\to\infty} \frac{1}{n} \log \| M_1 M_2 \ldots M_n \|,$$

существующему почти наверное. Норма здесь определяется как обычно:

$$\| M \| = \max_i \sum_j | m_{ij} |, \quad M = \{m_{ij}\}.$$

Кроме того, при более сильных ограничениях они вывели асимптотическое распределение элементов $\log m_{ij}^{(n)}$, $N_n = \{m_{ij}^{(n)}\} = M_1 M_2 \ldots M_n$. Хотя эти результаты весьма специальны, они заслуживают серьезного внимания как стимулирующие дальнейшую работу.

В этой связи заслуживает упоминания вероятностный функциональный анализ, скажем, изучение стохастических функциональных уравнений, их итерационных решений и т. п. Подобные проблемы изучались Шпачеком, Ганшем, Бхаруча-Ридом и некоторыми другими авторами.

Теперь, после того как мы сделали беглый обзор теории и мотивов ее развития, мы можем приступить к детальному изложению.

Глава 2

СТОХАСТИЧЕСКИЕ ПОЛУГРУППЫ

2.1. Общие замечания

Вероятностные меры, рассматриваемые в этой книге, будут иметь своей областью определения сепарабельное топологическое хаусдорфово пространство S (определение сепарабельности см. в замечаниях 2.1). Вероятностная мера P определяется на множествах, принадлежащих σ-алгебре $\mathscr{A}$, образованной всеми открытыми множествами из S (см. замечания 2.1.$_1$). Иногда мы будем также использовать неограниченные меры m ($m(S)$ не конечна); в некоторых из следующих далее утверждений это подразумевается.

Мы будем использовать терминологию, известную из теории вероятностей и теории конечномерных евклидовых пространств, часть понятий непосредственно переносится на наш случай. Два случайных элемента из S называются независимыми, если их совместное распределение на $S \times S$ есть произведение мер; аналогично определение независимости для нескольких случайных элементов. Распределение вероятностей P называется непрерывным, если $P\{s\} = 0$ для любого множества $\{s\}$, состоящего из единственного элемента s. Оно называется дискретным, если вся распределенная масса сосредоточена на элементах $s_1, s_2, \ldots$ с точечными массами $p_1, p_2, \ldots, p_n, \ldots \geqslant 0$, $\sum\limits_1^\infty p_n = 1$. Будем говорить, что P абсолютно непрерывно относительно заданной σ-конечной меры, если

$$P(E) = \int\limits_E p(s)\, m(ds) \quad \text{для всех} \quad E \in \mathscr{A};$$

$p(s)$ называется плотностью распределения P.

Большое внимание будет уделяться локально компактным пространствам. Мы рассматриваем борелевские меры, определенные на σ-алгебре, порожденной всеми компактными множествами. Удобно предполагать (и это не будет очень сильным ограничением), что вероятностная мера регулярна, т. е.

$P(E) = \inf P(O)$, где O — открытое множество, содержащее E,

и

$P(E) = \sup P(C)$, где C — компактное множество, содержащееся в E.

В настоящем контексте предположение регулярности выполнено автоматически, как указано в замечаниях $2.1._1$. Следует заметить, что в несепарабельном случае это уже не так.

Если распределенная в соответствии с нашей вероятностной мерой масса не выходит за пределы некоторой части пространства S, то мы можем рассматривать только эту часть, выбирая ее настолько малой, насколько это возможно. Для этого служит понятие *носителя* $s(P)$ меры P. Из возможных определений носителя мы выбираем следующее: $s(P)$ определяется как множество тех точек s, каждая окрестность которых имеет положительную меру. Носитель есть замкнутое подмножество пространства S.

Для меры m описанного типа (борелевской, регулярной) можно образовать интеграл

$$I(f) = \int_S f(s)\, m(ds)$$

по крайней мере для некоторых действительных функций $f(s)$. Он, конечно, определен для $f \in L(S)$, где $L(S)$ — множество всех непрерывных действительных функций, каждая из которых обращается в нуль вне некоторого компактного множества. $I(f)$ есть линейный функционал на $L(S)$; кроме того, это положительный функционал: $I(f) \geqslant 0$, если $f \in L^+(S)$, где $L^+(S)$ — множество неотрицательных функций, принадлежащих $L(S)$. Это утверждение может быть обращено. Если заданный функционал $I(f)$, определенный на $L(S)$, обладает этими свойствами, то существует однозначно определенная (регулярная) борелевская мера m, такая, что имеет место написанное выше интегральное представление (см. замечания $2.1._2$).

Теперь введем норму

$$\| f \| = \sup_{s \in S} | f(s) |$$

на множестве всех непрерывных комплекснозначных функций $f(s)$, обращающихся в нуль на бесконечности. Ограниченные линейные функционалы получающегося банахова пространства $L_\omega(S)$ могут быть представлены в виде

$$\mu(f) = \int_S f(s)\,\mu(ds),$$

где μ — (регулярная) комплексная мера, а норма $\|\mu\|$ функционала задается полной (абсолютной) вариацией меры μ (см. замечания 2.1.$_2$).

При отыскании предельных теорем нужно специфицировать тип сходимости распределений вероятности. Два главных типа — это L-слабая[1]) и слабая сходимости (терминология в литературе не установилась). Последовательность вероятностных мер P_1, P_2, P_3, ... на S называется *L-слабо сходящейся* к мере m, если

$$\int_S f(s)\,P_n(ds) \to \int_S f(s)\,m(ds) \tag{1}$$

при $n \to \infty$ для любой $f \in L(S)$. Однако мы будем чаще пользоваться следующим определением. Если (1) имеет место для любой ограниченной и непрерывной функции f, то мы говорим о *слабой сходимости*. (Из последней, очевидно, вытекает, что предельная мера m является вероятностной.) Эти два понятия совпадают, если S компактно. Вообще последовательность вероятностных мер P_n слабо сходится к предельной мере Q тогда и только тогда, когда она L-слабо сходится и $Q(S) = 1$. Необходимость условия очевидна. Для доказательства достаточности заметим, что если имеет место L-слабая сходимость $P_n \to Q$ и $Q(S) = 1$, то для любого $\varepsilon > 0$ найдется такое компактное множество C_ε, что $Q(C_\varepsilon) > 1 - \varepsilon$ и $P_n(C_\varepsilon) > 1 - \varepsilon$ для всех n. Но если $f(s)$ — произвольная ограниченная и непрерывная функция, то мы можем равномерно аппроксимировать ее

[1]) В подлиннике «*vague convergence*». Ввиду отсутствия соответствующего русского термина мы приняли указанный вариант.— *Прим. перев.*

на C_ε некоторой функцией $f_\varepsilon \in L(S)$. Следовательно,

$$\int_S f(s)\,P_n(ds) - \int_S f(s)\,Q(ds) = \int_S f_\varepsilon(s)\,[P_n(ds) - Q(ds)] + \int_{C_\varepsilon} [f(s) - f_\varepsilon(s)]\,[P_n(ds) - Q(ds)] + \int_{S-C_\varepsilon} [f(s) - f_\varepsilon(s)]\,[P_n(ds) - Q(ds)],$$

где каждое из трех слагаемых в правой части может быть сделано сколь угодно малым. Заметим также, что множество $\mathscr{P}(S)$ регулярных борелевских вероятностных мер на S компактно в L-слабой топологии (см. замечания 2.1.$_3$).

Далее мы будем обозначать через δ_s вырожденную вероятностную меру, при которой вся распределенная масса сосредоточена на элементе $s \in S$. В частности, сходимость по вероятности к фиксированному элементу принимает вид $P_n \to \delta_s$, $s \in S$.

2.2. Стохастические полугруппы

Пусть теперь пространство S имеет также алгебраическую структуру и предположим, что она образует *топологическую полугруппу*. На S определено умножение — бинарная операция, которая должна быть непрерывной функцией обоих аргументов. Она должна быть ассоциативна, но не обязательно коммутативна. Как и в предыдущем разделе, мы можем ввести (регулярно) борелевскую вероятностную меру на S, и нашей первой задачей является установление связи между вероятностными и алгебраическими свойствами S.

Пусть мы наблюдаем два случайных элемента s и t соответственно из популяций P и Q на локально компактной полугруппе S. Образуем новый случайный элемент $u = st$. Что можно сказать о его распределении вероятностей? Для определения этого распределения рассмотрим

$$R(f) = \int_S \int_S f(st)\,P(ds)\,Q(dt), \tag{1}$$

где интеграл берется относительно прямого произведения мер $P \times Q$. Этот функционал определен для всех $f \in L(S)$ и положителен.

Определение 2.2. *Композицией $P * Q = R$ называется однозначно определенная (регулярная) вероятностная мера R, удовлетворяющая соотношению*

$$R(f) = \int\limits_S f(s)\, R(ds).$$

Можно выразить композицию $P * Q$ и непосредственно

Теорема 2.2.1. *Композиция может быть записана в виде*

$$R(B) = P * Q(B) \doteq \int\limits_S \int\limits_S I_B(st)\, P(ds)\, Q(dt) = T(B) \qquad (2)$$

для любого борелевского множества B из S. Здесь $I_B(s)$ обозначает характеристическую функцию множества B.

Доказательство. Достаточно показать, что $R(C) = T(C)$ для любого компактного множества $C \subset S$, так как регулярные меры определены их значениями на компактных множествах. Покроем множество C открытым множеством O $(C \subset O)$, таким, что $R(O) < R(C) + \varepsilon$ и $T(O) < T(C) + \varepsilon$ для заданного положительного ε. Введем функцию

$$f(s) = \begin{cases} 1, & \text{если } s \in C, \\ \text{непрерывна для всех } s, \\ 0, & \text{если } s \overline{\in} O. \end{cases}$$

Тогда

$$R(C) \leqslant \int\limits_S f(s)\, R(ds) = \int\limits_S \int\limits_S f(st)\, P(ds)\, Q(dt) \leqslant$$

$$\leqslant \int\limits_S \int\limits_S I_O(st)\, P(ds)\, Q(dt) = T(O) < T(C) + \varepsilon.$$

Аналогично

$$T(C) = \int\limits_S \int\limits_S I_C(st)\, P(ds)\, Q(dt) \leqslant \int\limits_S \int\limits_S f(st)\, P(ds)\, Q(dt) = $$
$$= \int\limits_S f(s)\, R(ds) \leqslant R(O) < R(C) + \varepsilon,$$

и два полученных неравенства доказывают теорему (см. замечания $2.2._1$).

Замечание. Если S является также топологической *группой*, то возможны дальнейшие упрощения, так как очевидно, что

$$P * Q(B) = T(B) = \int\limits_S P(Bt^{-1})\, Q(dt) = \int\limits_S Q(s^{-1}B)\, P(ds).$$

Мы будем часто использовать обозначение $\mathscr{P}(S)$ для множества всех борелевских вероятностных мер на S. Не очень трудно показать, что $\mathscr{P}(S)$ есть топологическая полугруппа, в которой бинарной операцией является композиция (см. замечания $2.2._2$). Также можно показать, что $\mathscr{M}(S)$ (множество всех ограниченных борелевских комплексных мер) образует банахову алгебру (с обычным определением сложения и умножения на скаляр), где умножение определяется как композиция, а норма как полная абсолютная вариация.

Имеется полезное и простое выражение носителя композиции $s(P * Q)$.

Теорема 2.2.2. *Носитель композиции $P * Q$ есть замыкание произведения носителей P и Q*:

$$s(P * Q) = \overline{s(P) \cdot s(Q)}$$

(см. замечания $2.2._3$).

Доказательство. Для того чтобы доказать, что $s(P * Q) \subset \overline{s(P) \cdot s(Q)} = A$, заметим, что

$$P * Q(A) \geqslant \int\limits_S \int\limits_S I_{s(P)\, s(Q)}(st)\, P(ds)\, Q(dt) \geqslant$$
$$\geqslant \int\limits_S I_{s(P)}(s)\, P(ds) \int\limits_S I_{s(Q)}(t)\, Q(dt) = 1,$$

так что $P * Q[\overline{s(P)\, s(Q)}] = 1$ и $s(P * Q) \subset \overline{s(P)\, s(Q)}$.

Для того чтобы доказать, что $s(P*Q) \supset \overline{s(P)\,s(Q)}$, выберем произвольный элемент $s \in \overline{s(P)\,s(Q)}$ и некоторую его окрестность N. Можно найти элементы $s' \in s(P)$, $s'' \in s(Q)$ со столь малыми окрестностями N' и N'', что $N'N'' \subset N$. Но тогда $P*Q(N) \geqslant P*Q(N'N'') \geqslant$ $\geqslant P(N')\,Q(N'') > 0$, откуда вытекает, что $s \in s(P*Q)$, так что $\overline{s(P)\,s(Q)} \subset s(P*Q)$, как и утверждалось.

Заметим, что если S компактно, то множество $A = s(P)\,s(Q)$ автоматически замкнуто, и символ замыкания может быть опущен.

Однородные случайные процессы могут быть введены несколькими способами. Способ, избранный ниже, иногда удобен, но отличен от обычно используемого на действительной прямой (см. замечания 2.2.$_4$).

Определение 2.2.3. *Под однородным случайным процессом на полугруппе S понимается некоторое семейство вероятностных мер $\{P_t;\ 0 \leqslant t < \infty\}$ на S, таких, что $P_t * P_s = P_{t+s}$, $0 \leqslant t,\ s < \infty$. Этот процесс называется непрерывным (по вероятности), если $P_t \to \delta_e$ слабо при $t \downarrow 0$.*

Определим также *безгранично делимые распределения.*

Определение 2.2.4. *Распределение вероятностей $P \in \mathcal{P}(S)$ на полугруппе с единицей e называется безгранично делимым, если для любого натурального числа n существует $P_n \in \mathcal{P}(S)$, такое, что*

$$P_n^{n*} = P \ \text{и}\ P_n \to \delta_e \ \text{при}\ n \to \infty.$$

Замечание. Распределения P_t, образующие непрерывный однородный процесс, автоматически являются безгранично делимыми. Но нет оснований полагать, что заданное безгранично делимое распределение P может быть всегда вложено в однородный процесс P_t так, что $P_1 = P$.

Условие $P_n \to \delta_e$ существенно, но интересно посмотреть, что случится, если его опустить, особенно на полугруппах без единицы.

Заметим, что в определении используются одинаковые компоненты P_n. Если допускаются различные компоненты, то можно получить более широкое определение.

2.3. Компактные стохастические полугруппы

Для стохастических полугрупп в том общем виде, в каком они рассматривались в последнем разделе, сейчас не представляется возможным получить какие-нибудь существенные вероятностные результаты желаемого типа. Чтобы это сделать, нужно наложить на S дополнительные структурные ограничения (например, групповое свойство или еще что-нибудь). Этот подход будет развит в следующих главах, а сейчас мы рассмотрим, не входя в чрезмерные подробности, что имеет место в случае, когда S компактно.

Вспомним, что операция композиции $P_1 * P_2$ непрерывна по P_1 и P_2 и что используется слабая * сходимость (см. замечания $2.3._1$), так что $\mathcal{P}(S)$ компактно (см. замечания $2.1._3$). Компактная полугруппа имеет по крайней мере один идемпотент (см. замечания $2.3._2$), и мы ниже его конструктивно выделим. Компактность $\mathcal{P}(S)$ делает поиски предельных теорем более обещающими. Конечно, вполне может случиться, что данная последовательность $P, P^{2*}, P^{3*}, \ldots$ не сходится, например в случае циклической группы, в которой вся распределенная вероятностная масса сосредоточена в одном элементе, отличном от единицы. В этом случае «колебания» последовательности P^{n*} можно устранить либо выбором подпоследовательности, либо применением процедуры суммирования, описанной ниже.

Если $C(S)$ — множество непрерывных функций $f(s)$ на S с обычной нормой $\max |f(s)|$, то *вероятностный оператор*

$$Tf(s) = \int f(st)\, P(dt)$$

линеен, ограничен и $\|T\| = 1$. Введем осредненную сумму итераций

$$A_n = \frac{1}{n}(T + T^2 + \cdots + T^n),$$

которая является вероятностным оператором относительно вероятностной меры

$$\pi_n = \frac{1}{n}(P + P^{2*} + \cdots + P^{n*}).$$

Семейство функций $f(st)$ аргумента s и параметра t равностепенно непрерывно: для любых $\varepsilon > 0$ и $s \in S$ существует такая окрестность $N(s)$, что

$$|f(s't)-f(st)|<\varepsilon \text{ для всех } t\in S,$$

если $s'\in N(s)$. Это легко следует из компактности S. Но тогда семейство функций $A_n f(s)$ аргумента s и параметра n также равностепенно непрерывно. Тогда мы можем выделить сходящуюся подпоследовательность $A_{n_\gamma} f(s)$ (см. замечания 2.3.$_3$). Из одного варианта эргодической теоремы (см. замечания 2.3.$_4$) вытекает, что функции $A_n f(s)$ сходятся [по норме пространства $C(S)$] к некоторому $Af(s)$ и $AT = TA = A^2 = A$, причем A есть положительное линейное преобразование в $C(S)$. В переводе на вероятностный язык это означает следующее.

Теорема 2.3.1. *Среднее по Чезаро*

$$\pi = \lim_{n\to\infty}(1/n)(P + P^{2*} + \ldots + P^{n*})$$

существует и является идемпотентной мерой $\pi^{2*} = \pi$. *Далее*, $P * \pi = \pi * P = \pi$.

Можно немного более полно описать возможные предельные меры и идемпотенты. Прежде всего ясно, что при любом предельном распределении распределенная масса находится на замыкании суммы всех множеств $[s(P)]^n$, $n = 1, 2, \ldots$. Поэтому мы вполне можем ограничиться этим подмножеством множества S; очевидно, что оно тоже является компактной топологической полугруппой. В дальнейшем будет предполагаться, что это сделано. Можно показать затем, что $s(\pi)$ — ядро K полугруппы S (см. замечания 2.3.$_4$).

Пусть теперь P есть произвольная идемпотентная мера $P^{2*} = P$ с носителем $s(P) = \Sigma$. Это — подполугруппа, так как $\Sigma = \Sigma\cdot\Sigma$. Обозначим через K ее ядро. Покажем, что Σ есть вполне простая полугруппа (см. замечания 2.3.$_2$). Если $\Sigma = K$, то это ясно, так как ядро вполне простое. В противном случае $\Sigma - K$ непусто и должно содержать некоторый идемпотент, так как иначе мы получили бы противоречие: если I есть множество всех идемпотентов в Σ, то $\Sigma = \Sigma I \Sigma$ (см. замечания 2.3.$_2$), с другой стороны, все идемпотенты должны были бы содержаться в K, так что $I\subset K$ и $\Sigma I \Sigma \subset \Sigma K \Sigma \subset K$, откуда $\Sigma\subset K$, что противоречит допущению. Обозначим через i идемпотент в $\Sigma - K$ и пусть $f(s)$ есть непрерывная и неотрицательная функция на Σ, принимающая значения 1 на i и 0 на K. Рассмотрим функцию

$$f_1(s) = \int\limits_S f(ts)\, P(dt).$$

Пусть s_0 — точка, где $f_1(s)$ достигает максимума. Тогда

$$f_1(s_0) = \int\limits_S f(ts_0)\, P(dt) = \int\limits_S \int\limits_S f(\tau t s_0)\, P(d\tau)\, P(dt) = \\ = \int\limits_S f_1(ts_0)\, P(dt),$$

так что $f_1(s_0) = f_1(ts_0)$ для всех t из носителя $s(P) = \Sigma$. Но так как $Ks_0 \subset K$, то то же самое максимальное значение достигается в некоторой точке k ядра K. Однако

$$f_1(k) = \int\limits_S f(tk)\, P(dt) = 0,$$

так как $tk \in K$. Поэтому максимальное значение функции f_1 равно нулю. С другой стороны, $f_1(i)$ должно быть положительно, так как

$$f_1(i) = \int\limits_S f(ti)\, P(dt) \geqslant \int\limits_{N_i} f(ti)\, P(dt) \geqslant (1 - \varepsilon)\, P(N_i) > 0,$$

где N_i есть некоторая малая окрестность идемпотента i. Полученное противоречие доказывает, что $\Sigma = K$, и так как любое ядро является вполне простой полугруппой, отсюда вытекает следующая

Теорема 2.3.2.a. *Если P — идемпотентная вероятностная мера на компактной полугруппе, то ее носитель $s(P)$ — вполне простая подполугруппа.*

Можно следующим образом дать полную, хотя и не непосредственную, характеристику идемпотентных вероятностных мер на вполне простых полугруппах. Такие полугруппы могут быть представлены (см. замечания $2.3._2$) в виде

$$K = T \times X \times Y,$$

Здесь T есть компактная топологическая группа, а X и Y — компактные хаусдорфовы пространства. Умножение элементов

$$s=(t,\ x,\ y), \quad s'=(t',\ x',\ y')$$

определяется соотношением

$$ss'=(t\varphi(x,\ y')\,t',\ x',\ y),$$

где φ — непрерывная функция, принимающая значения из T и определенная на $X \times Y$. Тогда имеет место следующая теорема.

Т е о р е м а 2.3.2.б. *Пусть π—некоторая (регулярная) идемпотентная мера на компактной полугруппе. Носитель меры π является вполне простой подполугруппой K, и мера π может быть записана в виде*

$$\pi=\tau \times \xi \times \eta,$$

где τ — нормированная мера Хаара на T, а ξ и η — регулярные вероятностные меры соответственно на X и Y (см. замечания $2.3._2$).

Если полугруппа S еще и коммутативна, то можно получить более подробную информацию относительно идемпотентных мер. Это происходит потому, что в данном случае минимальный идеал образует группу, что, конечно, упрощает дело.

Т е о р е м а 2.3.3. *Если S — компактная коммутативная полугруппа и P — некоторая идемпотентная вероятностная мера, принадлежащая $\mathcal{P}(S)$, то носитель меры P есть подгруппа H полугруппы S, а P — нормированная мера Хаара на H.*

Для доказательства заметим, что нам уже известно, что носитель H совпадает с ядром K. Но K теперь является группой (см. замечания $2.3._5$), а мы увидим в гл. 3, теорема 3.2.1, что любая идемпотентная вероятностная мера, определенная на компактной группе, является нормированной мерой Хаара.

Теперь мы легко получаем следующий результат.

Теорема 2.3.4. *Для компактной и коммутативной стохастической полугруппы осредненные меры*

$$\pi_n = \frac{1}{n}(P + P^{2*} + \ldots + P^{n*})$$

сходятся к нормированной мере Хаара на наименьшем идеале K полугруппы, порожденной множеством $s(P)$

Доказательство. Мы уже знаем, что предельная мера π для последовательности π_n существует, является идемпотентной и имеет в качестве носителя минимальный идеал K. Теорема 2.3.3 утверждает, что π должна быть единственной нормированной мерой Хаара на K.

Рассмотрим теперь однородный процесс дискретного типа:

$$P_t * P_u = P_{t+u}; \quad t, \quad u \geqslant 0,$$

$$P_h(e) \to 1 \quad \text{при} \quad h \downarrow 0$$

на компактной стохастической полугруппе S с единицей e. Соответствующие вероятностные операторы

$$T_t f(s) = \int\limits_S f(su) P_t(du), \text{ где } f \in C(S),$$

образуют мультипликативную полугруппу. Так как $T_h - I$ есть оператор, соответствующий обобщенной мере (т. е. такой, которая может принимать и отрицательные значения) $M = P_h - \delta_e$, то его норма является абсолютной вариацией (см. 2.1):

$$\| T_h - I \| = \int\limits_S | M(ds) | = 1 - P_h(e) + P_h(\bar{e}) = 2[1 - P_h(e)] \to 0$$

при $h \downarrow 0$; полугруппа непрерывна в равномерной операторной топологии. Поэтому, как известно (см. замечания $2.3._6$), инфинитезимальный оператор существует:

$$V = \lim_{h \downarrow 0} \frac{T_h - I}{h}$$

и

$$T_t = \exp(tV).$$

Оператор V допускает представления

$$Vf(s) = \int\limits_S f(su) B(du),$$

где B — ограниченная обобщенная мера вариации 0, не имеющая отрицательной вариации нигде, кроме точки $s=e$. Возвращаясь от операторов к мерам, мы получаем следующую теорему.

Теорема 2.3.5. *Пусть P_t — однородный процесс дискретного типа на компактной стохастической полугруппе S. Тогда существует конечная обобщенная мера B с вариацией 0 и отрицательной вариацией только, быть может, в единице e, такая, что*

$$P_t = \exp^*(tB) = \sum_{k=0}^{\infty} t^k \frac{B^{k*}}{k!}.$$

Так как B может быть записано в виде $\lambda(Q-\delta_e)$, где $\lambda > 0$ и Q — вероятностная мера, то

$$P_t = \exp^*[t\lambda(Q-\delta_e)] = \exp(-t\lambda)\cdot\exp^*(t\lambda Q) = \sum_{k=0}^{\infty} \exp(-t\lambda)\frac{(t\lambda)^k}{k!} Q^{k*},$$

так как δ_e и Q коммутируют. Это дает возможность интерпретировать P_t как *сложное пуассоновское распределение*, образованное заданным распределением вероятностей Q и распределением Пуассона со средним значением λt. В нашу задачу не входит детальное исследование выборочных функций этого процесса, что потребовало бы введения сепарабельного варианта этого процесса, но ясно, что они должны вести себя, как в классическом случае прямой линии.

Вместо того чтобы требовать, чтобы $P_h \to \delta_e$ в указанном выше смысле, мы потребуем теперь, чтобы P_h сходилось к некоторой мере I:

$$\int_S |(P_h - I)(ds)| \to 0 \quad \text{при} \quad h \downarrow 0.$$

Тогда с помощью небольшой модификации предыдущего доказательства можно показать, что

$$P_t = I\exp^*(tQ),$$

где используются прежние обозначения, а I — идемпотентная мера и $Q = Q * I = I * Q$.

Будем предполагать теперь только, что $P_h \to \delta_e$ в смысле обычной слабой сходимости. Тогда T_t есть непрерывная полугруппа в сильной топологии и для любого $f \in C(S)$ имеет место сильная сходимость $T_h f \to f$ при $h \downarrow 0$. Поэтому из общей теории полугрупп вытекает, что имеет место сильная сходимость

$$T_t f = \lim_{\varepsilon \downarrow 0} \exp(tV_\varepsilon f),$$

где

$$V_\varepsilon = \frac{T_\varepsilon - I}{\varepsilon}.$$

Теорема 2.3.6. *Если P_t — непрерывный однородный процесс на компактной стохастической полугруппе, то P_t есть предел сложных пуассоновских процессов*:

$$P_t = \lim_{\varepsilon \downarrow 0} \exp^*(tB_\varepsilon),$$

где

$$B_\varepsilon = \frac{P_\varepsilon - \delta_e}{\varepsilon}.$$

Заметим, что если полугруппа S конечна, то операторы P_t [рассматриваемые на $C(S)$] также непрерывны в равномерной операторной топологии. Следовательно, инфинитезимальный порождающий оператор ограничен и определен на всем $C(S)$. Поэтому непрерывный однородный процесс на конечной полугруппе должен быть сложным пуассоновским процессом.

Однородные процессы относятся к безгранично делимым законам и поэтому уместно сделать относительно них несколько замечаний. Для начала предположим, что P есть такое распределение вероятностей, что для каждого натурального n существует $P_n \in \mathcal{P}(G)$, для которого

$$P = P_n^{n^*} \text{ и } \sup_n n[1 - P_n(e)] < \infty.$$

Для любого рационального $t = \nu/n$ можно определить однородный процесс P_t (где $P_1 = P$) так, как это будет сделано в разд. 3.2. Для того чтобы перейти к иррациональным значениям параметра t, достаточно (но не необходимо) показать, что $P_t(e) \to 1$ при $t \downarrow 0$ по всем рациональным значениям. А это так, поскольку

$$P_t(e) = P_n^{\nu *}(e) \geqslant [P_n(e)]^\nu \geqslant \left(1 - \frac{K}{n}\right)^\nu \to 1 \text{ при } \frac{\nu}{n} \to 0,$$

где $K = \sup_n n\,[1 - P_n(e)]$. Следовательно, заданное распределение P вкладывается в однородный процесс на стохастической полугруппе и можно непосредственно использовать то, что мы знаем о таких процессах.

В этой связи естественно доказать следующую простую предельную теорему.

Теорема 2.3.6.а. *Рассмотрим последовательность вероятностных мер $P_1, P_2, \ldots$, определенных на компактной топологической полугруппе с единицей $e \in S$, $P_n \in \mathcal{P}(S)$. Допустим, что с ростом n распределенная масса стремится сосредоточиться в e:*

$$P_n(e) = 1 - \frac{\lambda}{n} + o\left(\frac{1}{n}\right),$$

и что условное распределение вероятностей Q_n при условии $s \neq e$ удовлетворяет соотношению

$$\lim_{n\to\infty} \int_S |Q_n(ds) - Q(ds)| = 0$$

для некоторого $Q \in \mathcal{P}(S)$. Тогда P_n^{n*} слабо сходятся к сложному пуассоновскому распределению

$$\lim_{n\to\infty} P_n^{n*} = \exp^*[\lambda(Q - I)] = e^{-\lambda} \sum_{\nu=0}^{\infty} \frac{\lambda^\nu}{\nu!} Q^{\nu *}.$$

В частном случае, когда Q сконцентрировано в s, т. е. $Q = \delta_s$, этот предел является распределением Пуассона, соответствующим стохастическому элементу вида s_0^ν; ν есть случайная величина, распределенная по закону Пуассона со средним значением λ.

Доказательство. Рассмотрим вероятностный оператор

$$Tf(t) = \int_S f(ts)\,P(ds), \quad f \in C(S).$$

Имеем

$$Tf(t)=f(t)\left[1-\frac{\lambda}{n}+o\left(\frac{1}{n}\right)\right]+\frac{\lambda}{n}\int\limits_S f(ts)\,Q(ds)+$$
$$+\frac{\lambda}{n}\int\limits_S f(ts)\,[Q_n(ds)-Q(ds)],$$

так что

$$T=I+\frac{\lambda}{n}(Q-I)+\Delta_n,$$

где символ Q используется теперь также для обозначения вероятностного оператора, соответствующего вероятностной мере Q, и где $\|\Delta_n\|=o(1/n)$. Из того, что композициям P^{n*} соответствуют степени T^n и что

$$T^n f=\left[I+\frac{\lambda}{n}(Q-I)+\Delta_n\right]^n f\to\exp^*[\lambda(Q-I)]\,f,$$

следует утверждение теоремы.

Можно легко охарактеризовать носитель произвольного сложного пуассоновского распределения на компактной полугруппе.

Предложение. *Множество $E\subset S$ является носителем некоторого сложного пуассоновского распределения тогда и только тогда, когда E есть замкнутая подполугруппа группы S с единицей.*

Доказательство. Если P — сложное пуассоновское распределение:

$$P=\exp^*[\lambda(Q-\delta_e)]=e^{-\lambda}\sum_{\nu=0}^{\infty}\frac{\lambda^\nu}{\nu!}Q^{\nu*},$$

то $s(P)=\overline{\bigcup\limits_0^\infty s(Q^{\nu*})}$ при $\lambda>0$; случай $\lambda=0$ тривиален. P^{2*} имеет тот же вид, но с параметром 2λ, так что $s(P^{2*})=$ $=s(P)\,s(P)=s(P)$, откуда следует, что $s(P)$ есть подполугруппа с единицей. Будучи носителем, она автоматически замкнута. Обратно, если E обладает указанными свойствами, то мы можем определить распределение вероятностей Q с $s(Q)=E$ (можно выбрать счетное подмножество из Q, всюду плотное в Q, и приписать положительные вероятности каждому элементу этого подмножества). Образуем $P = \exp^* (Q - \delta_e)$. Сразу видно, что $s(Q^{\nu *}) = s(Q) = E$ и $s(P) = E$. Это предложение не имеет места для произвольного безгранично делимого распределения.

Сделаем теперь еще один шаг в направлении дальнейшей специализации и предположим, что S есть конечная и коммутативная стохастическая полугруппа с элементами $s_1, s_2, \ldots, s_n$. Тогда распределение вероятностей P на S характеризуется дискретными вероятностями $p_1, p_2, \ldots, p_n$; $\Sigma p_\nu = 1$. Будем также предполагать, что существует такое натуральное m, что $s^{m+1} = s$ для всех s (см. замечания 2.3.$_7$).

Хотя этот случай и может показаться слишком специальным, мы уделяем ему внимание, поскольку здесь мы впервые встречаемся с *гармоническим анализом* (*анализом Фурье*) — наиболее мощным аналитическим инструментом, используемым при изучении вероятностных предельных законов на общих структурах.

Определение 2.3.1. *Комплекснозначная функция $\sigma(s) \not\equiv 0$ называется полухарактером, если $\sigma(s)\,\sigma(t) = \sigma(st)$ для всех $s, t \in S$. Множество всех полухарактеров обозначается Σ.*

Число n элементов в Σ равно числу элементов в S. Множество Σ является полугруппой тогда и только тогда, когда полугруппа S имеет единицу, что мы и будем предполагать. Все n полухарактеров линейно независимы.

Определение 2.3.2. *Преобразованием Фурье $\hat{P}(\sigma)$ распределения P мы называем определенную на Σ функцию*

$$\hat{P}(\sigma) = \mathbf{E}\sigma(s) = \sum_{\nu=1}^{n} p_\nu \sigma(s_\nu).$$

Теорема 2.3.7. а) *Распределение P однозначно определяется преобразованием $\hat{P}$ по формуле обращения*

$$p_\nu = \sum_j \hat{P}(\sigma_j)\, A_{j\nu},$$

где $A_{j\nu}$ — коэффициенты, не зависящие от P (теорема единственности).

б) $\hat{P}_\nu \to \hat{P}$ *тогда и только тогда, когда* $P_\nu \to P$ (*теорема непрерывности*).

в) $\widehat{(P_1 * P_2)} = \hat{P}_1\hat{P}_2$: *композиции соответствует произведение преобразований Фурье.*

г) *Функция* $f(\sigma)$ *может быть представлена в виде преобразования Фурье* $\hat{P}$ *некоторого распределения вероятностей на* S *тогда и только тогда, когда она является положительно определенной:*

$$\sum_{\sigma,\, \psi \in \hat{S}} C_\sigma \overline{C}_\psi f(\sigma\overline{\psi}) \geqslant 0 \quad и \quad f(1) = 1,$$

где 1 *обозначает полухарактер, тождественно равный единице.*

Доказательство. а) Так как полухарактеры линейно независимы, то определитель $\det(\sigma_j(s_\nu))$ не равен нулю и система n линейных однородных уравнений

$$\hat{P}(\sigma_j) = \sum_1^n p_\nu \sigma_j(s_\nu)$$

может быть решена относительно p_ν.

б) В одну сторону это утверждение очевидно, в другую оно вытекает из формулы обращения.

в) $$\widehat{(P_1 * P_2)} = \sum_{\nu,\, \mu=1}^n p_\nu p_\mu \sigma(s_\nu s_\mu) =$$
$$= \sum_{\nu=1}^n p_\nu \sigma(s_\nu) \sum_{\mu=1}^n p_\mu \sigma(s_\mu) = \hat{P}_1(\sigma)\hat{P}_2(\sigma).$$

г) Если $f(\sigma)$ есть преобразование Фурье распределения P, то $f(1) = \sum_\nu p_\nu \cdot 1 = 1$ и

$$\sum_{\sigma,\, \psi} C_\sigma \overline{C}_\psi f(\sigma\overline{\psi}) = \sum_{\sigma,\, \nu} C_\sigma \overline{C}_\psi p_\nu \sigma(s_\nu)\overline{\psi}(s_\nu) =$$
$$= \sum_\nu p_\nu \Big| \sum_\sigma C_\sigma \sigma(s_\nu) \Big|^2 \geqslant 0. \qquad (1)$$

С другой стороны, если $f(\sigma)$ — произвольная комплекснозначная функция на Σ, то она может быть представлена в виде преобразования Фурье некоторой комплекснозначной (не обязательно неотрицательной) функции, определенной на S, скажем со значениями $s_\nu = q_\nu$. Выбирая C так, чтобы

$$\sum_\sigma C_\sigma \sigma(s_\nu) = \begin{cases} 1, & \text{если} \quad \nu = \mu, \\ 0, & \text{если} \quad \nu \neq \mu, \end{cases}$$

мы видим, что q_μ задается соотношением (1) и неотрицательно. Кроме того, $\sum_\nu p_\nu = f(1) = 1$.

Основной результат, относящийся к предельным законам на коммутативной конечной стохастической полугруппе, заключается в следующем.

Теорема 2.3.8. *Пусть i — идемпотент в S, такой, что $is^m(P) \subset S_i = \{s \mid s^m = i\}$. Обозначим через Σ_i полугруппу, порожденную $is^m(P)$. Последовательность P^{n*} имеет предел тогда и только тогда, когда $is(P)$ имеет общий элемент с Σ_i при любом i, определенном выше.*

Укажем только основную идею доказательства. Из того, что было сказано выше, ясно, что последовательность P^{n*} не сходится только тогда, когда $\hat{P}(\sigma) = e^{i\theta}$ для некоторого σ, где θ — действительное число, такое, что $\theta \not\equiv 0 \bmod 2\pi$. Но тогда

$$1 = |\hat{P}(\sigma)| \leqslant \sum_\nu p_\nu |\sigma(s_\nu)| = \sum_\nu p_\nu = 1,$$

и, следовательно, $\sigma(s_\nu) = e^{i\theta}$ для всех $s_\nu \in s(P)$. Можно убедиться (см. замечания $2.3._7$), что это эквивалентно условию теоремы. С таким же рассуждением мы встретимся позднее при изучении стохастических групп, и тогда мы сообщим бóльшие подробности, относящиеся к доказательствам.

Конечные стохастические полугруппы (коммутативные или нет) могут быть с успехом изучены с точки зрения конечных *марковских цепей*. Переходные вероятности задаются соотношением

$$p_{ij} = \sum_{s_i s_\nu = s_j} p_\nu,$$

где суммирование ведется по всем ν, для которых $s_i s_\nu = s_j$, и можно применить результаты теории цепей Маркова.

Мы знаем, например, что если существуют натуральное k и индекс j, такие, что $p_{ij}^{(k)} > 0$ для всех i, то последовательность P^{n*} сходится к предельному распределению. Если, кроме того, S — группа, то

$$\sum_{i=1}^{n} p_{ij} = \sum_{s_\nu = s_i^{-1} s_j} p_\nu = \sum_{\mu} p_\mu = 1,$$

где во второй сумме j фиксировано, а i пробегает все значения от 1 до n. Таким образом, матрица переходных вероятностей является дважды стохастической [1]). Предельное распределение является тогда просто равномерным: $P(s = s_\nu) = 1/n$. Аналогичным образом можно рассмотреть более интересные случаи с бóльшим числом нулей в матрице переходных вероятностей.

2.4. Примеры

Поясним сказанное небольшим числом простых примеров. Рассмотрим стохастическую полугруппу второго порядка со следующей таблицей умножения:

	1	2
1	1	2
2	2	2

Это — коммутативная полугруппа. Матрица переходных вероятностей имеет вид

$$P = \begin{Bmatrix} p_1 & p_2 \\ 0 & 1 \end{Bmatrix},$$

где $p_1 = P(1)$, $p_2 = P(2)$. Если $p_1 > 0$ и $p_2 > 0$, так что $s(P) = S$, то существует предельное распределение с $P(1) = 0$, $P(2) = 1$. Только в случае $p_1 = 1$, $p_2 = 0$

[1]) То есть, как сумма элементов, стоящих в одном столбце матрицы, так и сумма элементов, стоящих в одной ее строке, равна единице.— *Прим. перев.*

на элементе 1 в предельном распределении сосредоточена положительная масса, и тогда $P(1) = 1$.

В тех же обозначениях, но с таблицей умножения

	1	2
1	1	2
2	2	1

,

соответствующей коммутативной группе, матрица переходных вероятностей имеет вид

$$P = \begin{Bmatrix} p_1 & p_2 \\ p_2 & p_1 \end{Bmatrix}.$$

Если $p_1 > 0$ и $p_2 > 0$, то существует равномерное предельное распределение $P(1) = P(2)$ на носителе меры P. Если $p_1 = 0$, то предельного распределения не существует, но среднее по Чезаро мер P^{n*} является равномерным распределением. Если $p_1 = 1$, то существует равномерное предельное распределение $P(1) = 1$ на носителе меры P.

Пусть S — множество точек интервала $[0, 1]$ с бинарной операцией $st = \max(s, t)$. Для заданного распределения P на $[0, 1]$ пусть M есть верхняя грань $s \in S$ относительно этой меры P: $M = \sup s$; $P[0, s] < 1$. Конечно, достаточно ограничиться подполугруппой $[0, M]$. Все элементы S являются идемпотентами, и ядро K подполугруппы $[0, M]$ состоит из единственного элемента M. Из общих соображений ясно, что среднее по Чезаро мер P^{n*} имеет всю массу распределенной на ядре K; в этом специальном случае последовательность P^{n*} сходится к вырожденной предельной мере с $P(M) = 1$.

Пусть случайные величины $x_{n1}, x_{n2}, \ldots, x_{nn}$ независимы и имеют функцию распределения $F_n(x)$ на $[0, 1]$. Тогда (полугрупповое) произведение $x_{n1}x_{n2} \ldots x_{nn}$ имеет функцию распределения $F_n^n(x)$. Если предположить, что предел

$$G(x) = \lim_{n\to\infty} n[1 - F_n(x)]$$

существует на $[0, 1]$, то последовательность P_n^{n*} стремится к предельному распределению с функцией распределения

$$F(x)=\exp[-G(x)]$$

на $[0, 1]$. Функция $G(x)$ не возрастает, и $G(1)=0$, $G(0) \leqslant +\infty$. Ее можно интерпретировать как асимптотически ожидаемое число значений $x_{n1}, x_{n2}, \ldots, x_{nn}$, попадающих в интервал $(x, 1]$. (Отметьте связь с классической теорией «крайних значений» вариационного ряда.)

Какой вид имеют непрерывные однородные процессы P_t? Описывая P_t функцией распределения $F_t(x)$, мы видим, что, поскольку $F_t(x)$ удовлетворяет соотношению

$$F_{t+s}(x)=F_t(x)*F_s(x)=F_t(x)F_s(x),$$

она должна иметь вид

$$F_t(x)=\exp[-tG(x)],$$

где $G(x)$ есть функция, обладающая указанными выше свойствами. Легко найти инфинитезимальный порождающий оператор для $\{P_t\}$. Действительно, пусть $f(x) \in C'(S)$, где $C'(S)$ — множество непрерывных функций на $S = [0, 1]$, для которых интеграл

$$\int_0^1 |f(x)-f(0)|\,G(dx)$$

сходится. Последнее ограничение относится к поведению $f(x)$ только в точке $x=0$, поэтому $C'(S)$ всюду плотно в $C(S)$. Имеем

$$A_h f(x)=\frac{T_h-I}{h}f(x)=\frac{1}{h}\int_0^1 [f(\max(x, y))-f(x)]\,P_h(dy)=$$

$$=\frac{1}{h}\int_x^1 [f(y)-f(x)]\,F_h(dy)=$$

$$=-\int_x^1 [f(y)-f(x)]\exp[-hG(y)]\,G(dy).$$

При $h \to 0$ эта величина сходится к

$$Af(x) = -\int_x^1 [f(y) - f(x)]\, G(dy),$$

так что оператор A с областью определения $C'(S)$ является инфинитезимальным. Если, в частности, значение $G(0)$ конечно, то A — ограниченный оператор, P_t — сложный пуассоновский процесс и

$$F_t(x) = \sum_{\nu=0}^{\infty} \frac{(t\lambda)^\nu}{\nu!} e^{-\lambda t} H^\nu(x),$$

где $H(x) = \lambda - G(x)$, $\lambda = G(0)$.

Возникает вопрос о существовании нормальных процессов на S, т. е. существует ли непрерывный однородный процесс P_t с независимыми приращениями, почти все выборочные функции которого непрерывны. Мы имеем дело только с сепарабельными процессами. Рассмотрим непрерывный однородный процесс x_t. Пусть n — большое число. Тогда можно записать

$$x_{\nu/n} = \max[\eta_1^{(n)}, \eta_2^{(n)}, \ldots, \eta_\nu^{(n)}], \qquad 0 \leqslant \nu \leqslant n,$$

где случайные величины $\eta_\nu^{(n)}$ независимы и имеют общую функцию распределения

$$F_{1/n}(x) = \exp\left[-\frac{1}{n} G(x)\right].$$

Рассмотрим точки a и b, такие, что $0 < a < b < 1$ и $G(b) > 0$. Тогда вероятность того, что в последовательности $\eta_1^{(n)}, \eta_2^{(n)}, \ldots, \eta_n^{(n)}$ имеется $n-1$ значений из $[0, a]$ и одно значение из $(b, 1]$, равна

$$n \exp\left(-\frac{n-1}{n} G(a)\right)\left[1 - \exp\left(-\frac{1}{n} G(b)\right)\right].$$

Для больших значений n это приближенно равно

$$p = G(b) \exp[-G(a)] > 0.$$

Но отсюда вытекает, что с вероятностью $p - \varepsilon$ некоторое приращение $x_{(\nu+1)/n} - x_{\nu/n}$ больше, чем $b - a$. Следовательно, с вероятностью единица процесс не является непрерывным и не существует никакого нормального процесса на этой полугруппе.

Определяя $st = t$ для s и $t \in [0, 1]$, мы получим некоммутативную полугруппу. Все ее элементы по-прежнему являются идемпотентами; ядро совпадает со всей полугруппой. Так как $P^{n*} = P$, предельная мера всегда тривиальным образом существует и может иметь любой вид.

Другой, менее тривиальный пример дают матрицы второго порядка

$$\begin{Bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{Bmatrix}$$

с элементами $a_{ij} \geqslant 0$ и нормой $\| A \| \leqslant 1$. Норма определяется равенством

$$\| A \| = \max_i \sum_j | a_{ij} |$$

и обладает обычным свойством $\| AB \| \leqslant \| A \| \cdot \| B \|$. Это — компактная полугруппа. Если $s(P) = S$, то $\mathbf{E} \, \|A\| < 1$ (в действительности это имеет место уже тогда, когда на множестве $\| A \| < 1$ сосредоточена положительная масса), и для $B_n = A_1 A_2 \ldots A_n$ имеем

$$\mathbf{E} \| B_n \| \leqslant [\mathbf{E} \| A \|]^n \to 0$$

при $n \to \infty$, так что произведение B_n независимых элементов стохастической полугруппы сходится по вероятности к нулевой матрице. Этого и нужно ожидать, исходя из приведенной выше теории. Действительно, так как 0 есть идемпотентный элемент из S, то ясно, что минимальный идеал (ядро S) $K = \bigcap_i SiS$, где i — произвольный идемпотент, должен состоять из одного элемента 0.

Рассмотрим подмножество Σ матриц вида

$$A = \begin{Bmatrix} x & 1-x \\ y & 1-y \end{Bmatrix},$$

где $0 \leqslant x \leqslant y \leqslant 1$, оно является замкнутой подполугруппой исходной полугруппы S. Пусть при распределении P распределенная масса сосредоточена на подмножестве множества Σ, содержащем по крайней мере одну точку, для которой x или $y \neq 0$; 1. Вводя обозначения

$$A_\nu = \begin{Bmatrix} x_\nu & 1-x_\nu \\ y_\nu & 1-y_\nu \end{Bmatrix}, \quad B_n = A_1 \ldots A_n = \begin{Bmatrix} u_n & 1-u_n \\ v_n & 1-v_n \end{Bmatrix},$$

получаем

$$\begin{cases} u_{n+1} = y_{n+1} + u_n (x_{n+1} - y_{n+1}), \\ v_{n+1} = y_{n+1} + v_n (x_{n+1} - y_{n+1}). \end{cases}$$

Итерируя, нетрудно убедиться, что последовательность P^{n*} сходится к предельному распределению, задаваемому стохастической матрицей

$$C = \begin{Bmatrix} z & 1-z \\ z & 1-z \end{Bmatrix};$$

здесь z распределено как

$$z = y_1 + y_2 (x_1 - y_1) + y_3 (x_1 - y_1)(x_2 - y_2) + \ldots,$$

где (x_ν, y_ν) независимы для различных значений ν и распределены в соответствии с P. Этот ряд сходится с вероятностью единица, так как при сделанных предположениях $\mathbf{E}\,|x-y| < 1$, так что

$$\sum_1^\infty |\mathbf{E} y_{k+1} (x_1 - y_1)(x_2 - y_2) \ldots (x_k - y_k)| \leqslant \sum_1^\infty [\mathbf{E}\,|x-y|]^k < \infty.$$

Это вполне соответствует общей теории. Действительно, идемпотенты имеют вид

$$\begin{Bmatrix} 1 & 0 \\ 0 & 1 \end{Bmatrix} \quad \text{или} \quad \begin{Bmatrix} z & 1-z \\ z & 1-z \end{Bmatrix},$$

и, используя это, можно определить ядро, как и раньше, и получить после некоторых выкладок, что K совпадает со множеством всех матриц $\begin{Bmatrix} z & 1-z \\ z & 1-z \end{Bmatrix}$, где $0 \leqslant z \leqslant 1$.

Вообще говоря, существуют «более вырожденные» предельные распределения, при которых вся масса сосредоточена в идемпотентных элементах вида

$$\begin{Bmatrix} 1 & 0 \\ 1 & 0 \end{Bmatrix}, \quad \begin{Bmatrix} 0 & 1 \\ 0 & 1 \end{Bmatrix} \quad \text{или} \quad \begin{Bmatrix} 1 & 0 \\ 0 & 1 \end{Bmatrix}.$$

Заметим, что конечные или счетные стохастические полугруппы можно также изучать непосредственно, используя теорию марковских цепей с конечным или счетным множеством состояний.

Пусть $P_t = \exp^*[t(Q - \delta_e)]$ и $P'_t = \exp^*[t(Q' - \delta_e)]$ — два сложных пуассоновских процесса на компактной полугруппе S. Если S коммутативна, то очевидно, что $R_t = P_t * P'_t$ есть также сложный пуассоновский процесс $\exp^*[t(Q + Q' - 2\delta_e)]$. То же самое имеет место на произвольной компактной полугруппе, если Q и Q' (или P_t и P'_t) коммутируют.

Что будет в случае, когда S не обязательно коммутативна, но известно, что R_t есть сложный пуассоновский процесс? Полагая $R_t = \exp^*(tR)$, получаем простым вычислением, что $R = Q + Q' - 2\delta_e$.

Разлагая $\exp^*$ в степенной ряд и сравнивая члены второго порядка, получаем

$$(Q - \delta_e)*(Q' - \delta_e) = (Q' - \delta_e)*(Q - \delta_e).$$

При выборе $Q = \delta_s$, $Q' = \delta_{s'}$ последнее соотношение сводится к $ss' = s's$. Другими словами, если композиция двух любых сложных пуассоновских процессов всегда дает в результате сложный пуассоновский процесс, то полугруппа S должна быть коммутативной.

Учитывая роль, которую играют сложные пуассоновские процессы в предельных теоремах, мы видим, что при изучении предельных теорем для композиций неидентичных вероятностных мер мы не можем ожидать столь простого развития теории, как в случае действительной прямой. На действительной прямой предельные теоремы совершенно нечувствительны к тому, являются ли компоненты идентичными или нет. На некоммутативной полугруппе ситуация значительно более сложная.

Здесь мы коснулись указанной трудности в специальной обстановке, но то же явление встречается во всей теории некоммутативных стохастических структур. В данной книге мы не будем углубляться в детальное исследование случая неравных компонент.

Глава 3

СТОХАСТИЧЕСКИЕ ГРУППЫ; КОМПАКТНЫЙ И КОММУТАТИВНЫЙ СЛУЧАИ

3.1. Общие замечания о стохастических группах

В этой и двух следующих главах мы будем иметь дело с локально компактными и сепарабельными группами (см. замечания 2.1). Групповое свойство позволит получить более подробные результаты относительно вероятностных свойств этих структур.

Классический результат, относящийся к топологическим группам, в первоначальном виде восходящий к Хаару, утверждает, что существует нетривиальная *инвариантная* слева (регулярная) борелевская мера μ, такая, что для любого борелевского множества $E \subset G$ и любого элемента $g \in G$ имеет место равенство $\mu(gE) = \mu(E)$ и $\mu(O) > 0$ для любого непустого открытого множества O. Такая мера, мера Хаара, однозначно определена с точностью до умножения на действительную постоянную. Конечно, существует также аналогичная инвариантная справа мера ν.

Для произвольного фиксированного g функция множества $\mu(Eg)$ опять инвариантна слева и может быть записана в виде $\Delta(g)\cdot\mu(E)$. Легко видеть, что $\Delta(gh) = \Delta(g)\cdot\Delta(h)$ и что $\Delta(g)$ есть непрерывная положительная функция на G. Если группа G компактна, то $\Delta(g)$ должно равняться единице, так как если бы для некоторого элемента g_0 было $\Delta(g_0) \neq 1$, скажем > 1, то последовательность $\Delta(g_0^n) = \Delta^n(g_0)$ была бы неограничена, хотя $\Delta(x)$ должна быть ограничена в силу своей непрерывности и компактности своей области определения. Следовательно, меры, инвариантные слева, и меры, инвариантные справа, по существу, совпадают на компактной группе (на коммутативной группе это имеет место тривиальным образом), и мы можем их нормировать одинаковым образом: $\mu(G) = \nu(G) = 1$, поскольку компактные множества имеют конечные меры.

Множество $\mathcal{P}(G)$ всех (регулярных) нормированных борелевских мер на G, как и раньше, является топологической полугруппой. Можно было бы подумать, что $\mathcal{P}(G)$ теперь будет группой, но это не так, если только G не состоит из единственного элемента e. Действительно, пусть заданное распределение вероятностей $P \in \mathcal{P}(G)$ имеет обратное распределение Q, т. е. $P * Q = \delta_e$, где δ_g есть распределение вероятностей, при котором вся распределенная масса сосредоточена на элементе g. Но тогда $\overline{s(P)\, s(Q)} = e$, а это возможно, только если $s(P)$ состоит из единственной точки.

Важным подмножеством в $\mathcal{P}(G)$ является множество $\mathcal{A}(G)$ всех распределений из $\mathcal{P}(G)$, абсолютно непрерывных относительно μ (см. замечания 3.1.$_1$). Они могут быть записаны в виде

$$P(E) = \int_E p(g)\, \mu(dg),$$

где $p(g)$ — обобщенная *плотность*, или производная Радона — Никодима $dP/d\mu$. Если Q — произвольное распределение из $\mathcal{P}(G)$, то $Q * P \in \mathcal{A}(G)$ и

$$\frac{d(Q*P)}{d\mu} = \int_{h \in G} p(h^{-1}g)\, Q(dh),$$

так что $\mathcal{A}(G)$ является левым идеалом. Если Q также принадлежит $\mathcal{A}(G)$ и $dQ/d\mu = q$, то

$$\frac{d(Q*P)}{d\mu} = \int_{h \in G} p(h^{-1}g)\, q(h)\, \mu(dh).$$

С $\mathcal{A}(G)$ связана групповая алгебра $L_1(G)$, состоящая из всех комплекснозначных функций, абсолютно интегрируемых относительно меры Хаара на G. Если умножение в $L_1(G)$ определено, как и выше, интегралом композиции, то ясно, что $L_1(G)$ есть банахова алгебра (см. замечания ния 3.1.$_2$).

Аналогично определяется гильбертово пространство $L_2(G)$ как множество всех комплекснозначных функций, квадрат которых интегрируем относительно ν, с обычным в L_2 определением скалярного произведения.

При $f \in L(G)$ вероятностный оператор T, соответствующий распределению P на G, определяется формулой

$$Tf(g) = \mathbf{E}f(gh) = \int_{h \in G} f(gh)\, P(dh).$$

Область определения T может быть расширена до $L_2(G)$. Действительно, при норме пространства $L_2(G)$ имеем (см. замечания 3.1.$_3$)

$$\|Tf\|^2 \leqslant \int_G \int_G \int_G |f(gh_1) f(gh_2)|\, P(dh_1)\, P(dh_2)\, \nu(dg) \leqslant$$

$$\leqslant \int_G \int_G \sqrt{\int_G |f(gh_1)|^2\, \nu(dg)} \sqrt{\int_G |f(gh_2)|^2\, \nu(dg)}\, P(dh_1) P(dh_2)$$

и

$$\int_G |f(gh)|^2\, \nu(dg) = \int_G |f(u)|^2\, \nu(du\, h^{-1}) =$$

$$= \int_G |f(u)|^2\, \nu(du) = \|f\|^2,$$

так что $\|Tf\| \leqslant \|f\|$, $f \in L$, что обеспечивает единственное продолжение на $L_2(G)$ обычным образом, так как $L(G)$ плотно в $L_2(G)$.

Особый интерес представляет оператор

$$R_h f(g) = f(gh),$$

соответствующий вероятностной мере $P = \delta_h$. Оператор R_h называется *правым переносом*. Левый перенос может быть определен как

$$L_h f(g) = f(h^{-1}g).$$

Оператор, сопряженный с T, имеет вид

$$T^* f(g) = \int_{h \in G} f(gh)\, Q(dh),$$

где $Q(E) = P(E^{-1})$, так что $T^* f(g) = \mathbf{E}f(gh^{-1})$ при тех же обозначениях, что и выше. Действительно, $(Tf, \varphi) =$

$$= \int_G \int_G f(gh)\, \overline{\varphi(g)}\, P(dh)\, \nu(dg) = \int_G \int_G f(u)\, \overline{\varphi(uv)}\, Q(dv)\, \nu(du) =$$

$$= (f, T^*\varphi).$$

Необходимым и достаточным условием самосопряженности оператора T являются равенства $P(E) = Q(E) = P(E^{-1})$, т. е. мера P должна быть *симметричной*. Для того чтобы оператор T был нормальным, необходимо и достаточно, чтобы $P * Q = Q * P$; это доказывается аналогично.

Особенно важен симметричный случай. Спектр оператора T расположен на интервале $(-1, 1)$, и интересно рассмотреть случаи, когда $\| T \| = 1$ и когда $\| T \| < 1$. В последнем случае $T^n \to 0$ при $n \to \infty$, так что $P^{n*}(C) \to 0$ для любого компактного множества C. Это не может иметь места, конечно, когда G компактна. Предположим, что для любого C можно найти последовательность $\varphi_n(g) \in L(G)$, такую, что функция, тождественно равная единице, равномерно аппроксимируется на C функциями

$$\psi_n^{(h)} = \frac{\varphi_n * \widetilde{\varphi}_n}{\| \varphi \|^2} = \frac{1}{\| \varphi \|^2} \int\limits_G \varphi_n(g) \overline{\varphi}_n(gh) \, \nu(dg),$$

где $\widetilde{\varphi}_n(g) = \overline{\varphi}_n(g^{-1})$. (С аналогичным условием мы встретимся позднее, в гл. 5.) Тогда

$$\frac{\| T\varphi_n \|^2}{\| \varphi_n \|^2} = \frac{1}{\| \varphi_n \|^2} \int\int\int \varphi_n(gh_1) \overline{\varphi}_n(gh_2) P(dh_1) P(dh_2) \nu(dg) =$$

$$= \int\int \psi_n(h_1^{-1} h_2) P(dh_1) P(dh_2),$$

а это стремится к 1 при $n \to \infty$, так что $\| T \|$ должна быть равна 1.

Сделаем некоторые предварительные замечания относительно T. Записывая T в спектральном представлении как самосопряженный оператор

$$T = \int\limits_{-1}^{1} \lambda \, dE(\lambda),$$

где $E(\lambda)$ — разложение единицы, получаем для преобразования, соответствующего P^{2n*}, что

$$T^{2n} f = \int\limits_{-1}^{1} \lambda^{2n} \, dE(\lambda) f \to (E_1 + E_{-1}) f = Sf$$

при $n \to \infty$. Здесь E_1 и E_{-1} соответственно скачки функции $E(\lambda)$ в точках $\lambda = \pm 1$. Предельная для P^{2n*} мера (которая a priori не обязательно нормирована) должна быть идемпотентной, так как $S^2 = S$. Если $\lambda = \pm 1$ не являются собственными значениями, то $T^n f \to 0$, и можно снова сделать вывод, что вся распределенная масса убегает в бесконечность.

Пусть P — распределение, при котором масса содержится в окрестности единичного элемента $e \in s(P)$. Если T имеет собственное значение 1, то существует функция $f \in L_2(G)$, такая, что

$$f(g) = \int f(gh)\, P(dh).$$

Для того чтобы это было возможно, все функции $f(gh)$, где $h \in s(P)$, должны быть равны для почти всех g. Читателю может быть полезно представлять себе эти функции геометрически как векторы в $L_2(G)$. Вводя множество

$$H = \{h \mid f(gh) = f(g) \text{ для почти всех } g\},$$

мы видим, что $s(P)$ должен содержаться в H. Но очевидно, что H является замкнутой подгруппой группы G и должно поэтому совпадать с G. Как обычно, мы предполагаем, что G уже выбрана так, что она является наименьшей замкнутой подгруппой (исходной группы), содержащей $s(P)$. Тогда собственная функция постоянна почти всюду на G, а такая функция может принадлежать $L_2(G)$, только если G компактна. С другой стороны, если G компактна, то $f(g) \equiv 1$ есть собственная функция, соответствующая собственному значению $\lambda = 1$.

Это связано с тем, что было сказано выше относительно симметричных распределений. Действительно, если P есть симметричная мера, то последовательность $T^{2n}f$ всегда сходится, так что последовательность P^{2n*} слабо сходится к некоторой мере P^∞. Предельная мера не обязательно нормирована; она может даже обращаться в нуль [$P^\infty(E) = 0$ для любого компактного множества E]. Это зависит от существования возможных собственных значений $\lambda = \pm 1$ оператора T, т. е. $\lambda = 1$ для оператора T^2. Но T^2 соответствует P^{2*} — мере, носитель которой всегда содержит единичный элемент e, так как P предполагается симметричной.

Докажем теперь следующий результат, имеющий место при стандартных условиях (сепарабельная, локально компактная группа) без предположения симметричности P.

Теорема 3.0. *Образуем для заданной вероятностной меры $P \in \mathcal{P}(G)$ средние π_n итераций $P^{\nu *}$:*

$$\pi_n = \frac{1}{n} \sum_{\nu=1}^{n} P^{\nu *}.$$

Тогда π_n L-слабо[1] *сходится к некоторой предельной мере π, для которой $\pi(G) \leqslant 1$. При этом π есть идемпотентная вероятностная мера, т. е. $\pi^{2*} = \pi$, и либо $\pi(G) = 1$, либо π нулевая мера ($\pi = 0$).*

Доказательство. Имея вероятностный оператор

$$Tf(g) = \int\limits_G f(gh)\, P(dh),$$

мы можем использовать эргодическую теорему (см. замечания $3.1._4$) для доказательства того, что

$$f_n = T_n f = \frac{1}{n} \sum_{\nu=1}^{n} T^{\nu} f = f_n$$

сходится при $n \to \infty$. Достаточно считать, что $f \in L(G)$. Это и будет предполагаться ниже.

Мы утверждаем, что для любого положительного ε существует окрестность N_ε элемента e, такая, что

$$|f(g'h) - f(gh)| \leqslant \varepsilon \text{ для всех } h \in G,$$

если $g'g^{-1} \in N_\varepsilon$. Для того чтобы это доказать, обозначим через C компактное множество, вне которого $f(g)$ равно нулю, и выберем некоторую окрестность M элемента e. Множество $\overline{M}C$ компактно, и можно найти такую окрестность N_ε, что $N_\varepsilon \subset M$, $N_\varepsilon^{-1} \subset M$ и

$$|f(x') - f(x)| < \varepsilon,$$

если $x' \in \overline{M}C$, $x \in \overline{M}C$ и $x'x^{-1} \in N_\varepsilon$. Предположим, что $g'g^{-1} \in N_\varepsilon$, и рассмотрим четыре случая. 1) Если $g'h$

[1] См. стр. 39.— *Прим. ред.*

и $gh \in \overline{M}C$ и

$$d = | f(g'h) - f(gh) |,$$

то $d \leqslant \varepsilon$. 2) Если $g'h$ и gh лежат вне $\overline{M}C$, то $f(g'h) = f(gh) = 0$ и $d = 0$. 3) Если $gh \bar{\in} \overline{M}C$, то $g'h \bar{\in} C$, так как в противном случае $gh = g(g')^{-1}g'h \in MC$, что противоречит допущению; следовательно, $d = 0$. 4) Если $g'h \bar{\in} \overline{M}C$, то, аналогично, $d = 0$.

Равностепенная непрерывность, доказанная для функций $f(gh)$, где $h \in G$, имеет место также для функций $T^{\nu}f$ и функций

$$T_n f = \frac{1}{n} \sum_{\nu=1}^{n} T^{\nu} f = f_n.$$

Функции $f_n(g)$ сходятся почти наверное при $n \to \infty$ к некоторой функции $\varphi(g)$. В силу равностепенной непрерывности $f_n(e) \to \varphi(e)$, так что

$$\lim_{n \to \infty} \int_G f(h)\, \pi_n(dh) = \varphi(e).$$

Но $\varphi(e) = Sf$ есть положительный линейный функционал на $L(G)$ и может быть записан в виде

$$Sf = \int_G f(h)\, \pi(dh),$$

где π — мера, такая, что $\pi(G) \leqslant 1$. Она связана с оператором R, скажем в $L_2(G)$, формулой

$$Rf(g) = \int_G f(gh)\, \pi(dh).$$

Оператор R идемпотентен: $R^2 = R$. Если $\| R \| < 1$, то из равенства $R(R - I) = 0$ вытекает $R = 0$ и $\pi(G) = 0$. С другой стороны, если $\| R \| = 1$, то должно быть $\pi(G) = 1$. В противном случае, при $\pi(G) = k$ $(0 < k < 1)$, оператор R/k должен был бы быть вероятностным оператором. Такой оператор имеет норму $\leqslant 1$, что противоречит допущению. Доказательство теоремы закончено.

Из этой теоремы вытекает

Следствие. *Если носитель меры P содержится в компактной подполугруппе группы G, то $\pi(G) = 1$ и π_n слабо сходится к π. Если $\pi(G) = 1$, то носитель меры P содержится в компактной подгруппе группы G.*

Доказательство первой части очевидно. Далее, если $\pi(G) = 1$, то должна существовать такая функция $f \in L_2(G)$, что $f = Tf$. Рассуждая, как выше, можно показать, что $s(P)$ содержится в компактной группе.

Читатель, которого удивит соотношение между двумя утверждениями следствия, сможет лучше понять их связь, если заметит, что компактная подполугруппа группы образует подгруппу, так как в ней имеет место закон сокращения.

В классической теории суммирования независимых случайных величин основным орудием является анализ Фурье. Чтобы распространить этот метод на стохастические группы, можно попытаться ввести преобразование

$$\hat{P}(r) = \int_G r(g)\, P(dg), \quad r \in R,$$

где R — множество функций $r(g)$ на G, принимающих значения из некоторого, пока не определяемого, пространства V. Хотелось бы, чтобы, как и в классическом случае, имело место соотношение

$$\widehat{P_1 * P_2}(r) = \hat{P}_1(r) \cdot \hat{P}_2(r).$$

Ясно, что в V должно быть определено сложение, умножение на скаляр и умножение. Полагая $P_1 = \delta_{g_1}$, $P_2 = \delta_{g_2}$, получаем

$$r(g_1 g_2) = r(g_1) \cdot r(g_2),$$

а это приводит нас к рассмотрению группового представления G. Нужно также, чтобы пространство V было достаточно широким, для того чтобы имело место однозначное соответствие между распределениями и их преобразованиями, $P \longleftrightarrow \hat{P}$. Оказывается, что в коммутативном случае можно выбрать в качестве V множество комплексных чисел $e^{i\theta}$, по модулю равных единице; в компактном случае следует в качестве V выбрать множество всех конечномерных унитарных матриц, в то время как в случае произвольной локально компактной группы нужно использовать множество унитарных преобразований гильбертова пространства. Последний случай значительно сложнее первых двух и совсем не так хорошо изучен. Из дидактических соображений мы сначала изучим в ближайших двух разделах компактные и коммутативные группы, а исследование общих локально компактных групп будет отложено до гл. 5.

Прежде чем обратиться к анализу Фурье, рассмотрим кратко следующую проблему. Пусть h есть гомоморфное отображение локально компактной группы G на локально компактную группу G' (см. замечания 3.1.$_5$).

Теорема 3.1. *Распределение вероятностей $P \in \mathcal{P}(G)$ индуцирует распределение $P' \in \mathcal{P}(G')$ посредством соотношения*

$$\int_{G'} f(g')\,P'(dg') = \int_{G} f[h(g)]\,P(dg)$$

для любой функции $f \in L(G')$. Если $P_n \in \mathcal{P}(G)$ и P_n слабо сходится к P, то P'_n слабо сходится к P'. Если P_t — однородный процесс на G, такой, что $\lim\limits_{t\downarrow 0} P_t(NK) = 1$, где N — произвольная окрестность элемента e, а K — ядро гомоморфизма, то P'_t есть непрерывный однородный процесс на G'.

Доказательство. Рассмотрим линейный функционал

$$I(f) = \int_{G} f[h(g)]\,P(dg).$$

Заметим, что h должно быть открытым непрерывным отображением. Если $f(g')$ ограничена и непрерывна на G', то $f[h(g)]$ обладает теми же свойствами на G, и $I(f)$ однозначно определяет регулярную вероятностную меру P' на G', такую, что $I(f) = \int\limits_{G'} f(g')\,P'(dg')$. Для любого открытого множества $O \subset G$ множество $O' = h(O) \subset G'$ открыто и $P(O) = P'(O')$.

Пусть теперь P_n слабо сходится к P. Отсюда непосредственно вытекает слабая сходимость $P'_n \to P'$. На самом деле достаточно было бы предположить, что P_n сходится слабо к P по модулю K; читатель легко придаст этому утверждению точное значение и без труда его докажет.

Теперь пусть P_t — однородный процесс. Для любой ограниченной и непрерывной функции $f(g')$ имеем

$$\int_{G'} f(g') P'_{t+s}(dg') = \int_G f[h(g)] P_{t+s}(dg) =$$
$$= \int_G \int_G f[h(xy)] P_t(dx) P_s(dy) =$$
$$= \int_G \int_G f[h(x) h(y)] P_t(dx) P_s(dy) =$$
$$= \int_{G'} \int_{G'} f(x'y') P_t(dx') P_s(dy'),$$

откуда вытекает, что $P'_{t+s} = P'_t * P'_s$, так что P'_t — однородный процесс. Для того чтобы убедиться в том, что P'_t также слабо непрерывен, предположим, что f непрерывна и ограничена, и учтем, что

$$\int_{G'} f(g') P'_t(dg') = \int_G f[h(g)] P_t(dg) \to f(e') = \int_{G'} f(g') \delta_{e'}(dg'),$$

так как вероятностная масса P_t предполагается сходящейся к ядру K гомоморфизма и $h(g) = e'$ при $g \in K$. Следовательно, P'_t слабо сходится к $\delta_{e'}$, что и утверждалось.

Если группе G придана вероятностная структура с распределениями, однородными процессами и предельными законами, то из теоремы 3.1 вытекает ее аналог для фактор-группы $F = G/N$, где N есть замкнутая нормальная подгруппа (нормальный делитель). Надо только использовать естественный гомоморфизм G на F.

Надо сказать несколько слов об определении однородных случайных процессов на группах. Чтобы прийти к определению, более соответствующему обычному для случая прямой, чем определение 2.2.3, можно было бы сказать, что однородный процесс на группе — это случайный процесс $g(t)$, $0 \leqslant t < \infty$, принимающий значения из G и такой, что случайные элементы $g^{-1}(t_\nu)\,g(t_{\nu+1})$, $0 < t_1 < t_2 < \ldots$, независимы (это может быть названо левой инвариантностью и, конечно, можно было бы использовать также правую инвариантность) и распределение вероятностей $g^{-1}(t)\,g(s)$ зависит только от разности $s - t$. Если, кроме того, почти все выборочные функции непрерывны, естественно говорить о *броуновском движении* на группе. Изучение отдельных выборочных функций, однако, не входит в задачи этой книги.

3.2. Компактные стохастические группы

Изучим сначала идемпотентные меры на компактной группе G. Они полностью характеризуются следующей теоремой.

Теорема 3.2.1. *Если $P \in \mathcal{P}(G)$ есть идемпотентная мера на компактной группе G, то $s(P)$ — замкнутая подгруппа H группы G, а P — нормированная мера Хаара на H.*

Доказательство. Рассмотрим рассуждение, использованное в разд. 2.3 для доказательства соответствующей теоремы. Можно применить ту же идею, учитывая, что теперь $H = s(P)$ удовлетворяет условию $H^2 = H$, так что H является подгруппой (см. замечания 2.3.$_2$), Тогда идеал I равен H и функция g достигает своей верхней грани на всем H, т. е. g постоянна, и P есть инвариантная мера на H.

Перейдем теперь к анализу Фурье распределений вероятностей на компактных группах. Прежде всего вспомним основные факты, относящиеся к унитарным представлениям таких групп. Под унитарным представлением степени n мы понимаем функцию $M(g)$, значениями которой являются унитарные квадратные матрицы порядка n, $M(g) = \{m_{ij}(g);\ i, j = 1, 2, \ldots, n\}$, причем $m_{ij}(g)$ — непрерывные функции на G и функция $M(g)$ удовлетворяет основному соотношению $M(g_1)\,M(g_2) = M(g_1 g_2)$ для всех $g_1, g_2 \in G$. Такое представление называется неприводимым, если не существует подпространства унитарного n-мерного пространства, инвариантного слева относительно $M(g)$ для всех $g \in G$. Вместо представлений правильнее было бы говорить о классах эквивалентных представлений: два представления $M_1(g)$ и $M_2(g)$ называются эквивалентными, если $M_1(g) = AM_2(g)A^{-1}$, $g \in G$. Из каждого класса эквивалентности мы выберем по одному представителю.

Если r пробегает множество R всех неприводимых неэквивалентных унитарных представлений, то различные компоненты $m_{ij}^{(r)}(g)$ образуют множество функций на G, обладающее интересными свойствами. В гильбертовом пространстве $L_2(G)$ имеет место соотношение ортогональности: $m_{ij}^{(r)}(g) \perp m_{kl}^{s}(g)$ при $r \neq s$. Далее, n^2 элементов матрицы $M^{(r)}(g)$ ортогональны друг другу и имеют норму $n^{-1/2}$. Для сепарабельных групп, с которыми мы имеем дело, пространство $L_2(G)$ сепарабельно и, следовательно, R счетно, так что его элементы могут быть занумерованы индексами $r = 0, 1, 2, \ldots$, где индексом $r = 0$ отмечается тождественное представление $M(g) \equiv I$.

Основной результат, теорема Петера — Вейля, утверждает, что множество всех функций $m_{ij}^{(r)}(g)$ полно в $L_2(G)$. Оно даже равномерно полно, т. е. каждая непрерывная функция на G может быть равномерно аппроксимирована конечными линейными комбинациями функций $m_{ij}^{(r)}(g)$. Вследствие этого для каждого $g \neq e$ существует такое r, что $M^{(r)}(g) \neq I$, или, иными словами, неприводимых представлений достаточно много, чтобы разделить элементы группы.

Это естественным образом приводит к следующему определению.

Определение 3.2.1. *Под преобразованием Фурье вероятностной меры P на компактной группе G понимается последовательность матриц*

$$\hat{P}_r = \mathbf{E}M^{(r)}(g) = \int_G M^{(r)}(g)(P)(dg), \quad r = 0, 1, 2, \ldots .$$

Интегрирование по G здесь интерпретируется, конечно, как интегрирование матрицы по элементам. Перейдем теперь к изучению основных свойств таких преобразований Фурье.

Теорема 3.2.2. *Преобразование Фурье распределения вероятностей на компактной группе обладает следующими свойствами.*

а) *Распределение вероятностей* P *однозначно определяется преобразованием* $\hat{P}_r$, $r = 0, 1, 2, \ldots$

б) *Пусть* $P, P^{(1)}, P^{(2)}, \ldots$ — *последовательность распределений вероятностей и* $\hat{P}_r, \hat{P}_r^{(1)}, \hat{P}_r^{(2)} \ldots$ — *их преобразования Фурье. Тогда* $P^{(n)} \to P$ *эквивалентно* $\hat{P}_r^{(n)} \to \hat{P}_r$ *для всех* $r = 0, 1, 2, \ldots$

в) *Композиции двух вероятностных мер соответствует умножение преобразований Фурье,* $\widehat{P^{(1)} * P^{(2)}} = \hat{P}^{(1)} \cdot \hat{P}^{(2)}$.

г) *При любом* r *матрица* $\hat{P}_r$ *представляет ограниченное линейное преобразование,* $\| \hat{P}_r \| \leqslant 1$. *Для тождественного преобразования,* $r = 0$, *имеем* $\hat{P}_0 = I$.

Доказательство. Заданную непрерывную функцию $f(g)$ на G можно равномерно аппроксимировать конечными линейными комбинациями элементов $m_{ij}^{(r)}(g)$. Если $\hat{P}^{(1)} = \hat{P}^{(2)}$, то

$$\int_G f(g) P_1(dg) = \int_G f(g) P_2(dg)$$

для любой непрерывной функции $f(g)$, что доказывает а).

Точно так же из $\hat{P}^{(n)} \to \hat{P}$ следует, что для каждой непрерывной функции $f(g)$

$$\lim_{n\to\infty} \int_G f(g) P_n(dg) = \int_G f(g) P(dg),$$

т. е. $P_n \to P$. То, что из $P_n \to P$ вытекает $\hat{P}_n \to \hat{P}$, очевидно. Отображение $P \to \hat{P}$ является гомеоморфизмом.

Для доказательства в) достаточно заметить, что

$$\widehat{P^{(1)}*P^{(2)}}=\int\limits_G\int\limits_G M(g_1g_2)\,P^{(1)}(dg_1)\,P^{(2)}(dg_2)=$$
$$=\int\limits_G M(g_1)\,P^{(1)}(dg_1)\int\limits_G M(g_2)\,P^{(2)}(dg_2)=\hat{P}^{(1)}\cdot\hat{P}^{(2)}.$$

Отображение $P\to\hat{P}$ является изоморфизмом.

Последнее утверждение теоремы, г), вытекает из равенства $\|M_r(g)\|=1$, $g\in G$, так как P — вероятностная мера. Для $r=0$ представление $M_r(g)$ есть просто тождественное представление I, и $\hat{P}_0=\mathbf{E}I=I$.

Для заданного множества матриц A_0, A_1, A_2, ... было бы интересно знать, найдется ли такое $P\in\mathcal{P}(G)$, что $\hat{P}_r=A_r$, $r=0, 1, 2, \ldots$, и построить в этом случае P по A (обращение преобразования Фурье). В простейшем случае, если ряд

$$p(g)=\sum_{r,\,i,\,j}n_r a_{ij}^{(r)}\overline{m_{ij}^{(r)}}(g);\quad \{a_{ij}^{(r)}\}=A_r;$$

сходится равномерно к неотрицательной функции и если $A_0=I$, то A_r есть преобразование Фурье абсолютно непрерывного распределения вероятностей с плотностью $p(g)$. Чтобы убедиться в этом, заметим, что $p(g)$ непрерывна и, следовательно, интегрируема на G, $p\in L_1(G)$. Далее, в силу соотношений ортогональности для m,

$$\int\limits_G m_{ij}^{(r)}(g)\,p(g)\,\mu(dg)=a_{ij}^{(r)},$$

и, в частности, для $r=0$

$$\int\limits_G p(g)\,\mu(dg)=1,$$

что доказывает утверждение. В общем случае нужно применить какой-нибудь метод суммирования для обращения. Один из таких методов заключается в следующем. Рассмотрим убывающую последовательность окрестностей $N_1\supset N_2\supset N_3\supset\ldots$ единичного элемента, стягивающуюся к e: $\bigcap\limits_k N_k=e$. Для каждой N из этой последовательности обозначим через $f^N(g)$ неотрицательную функцию, равную нулю вне N и такую, что

$$\int_G f^N(g)\,\mu(dg) = 1.$$

Эти функции могут быть выбраны с большой свободой, но так, чтобы ряды

$$\sum_{r,\,i,\,j} n_r f^N_{rij}, \quad \text{где} \quad f^N_{rij} = \int_G f^N(g)\, m^{(r)}_{ij}(g)\,\mu(dg),$$

сходились абсолютно.

Теорема 3.2.3. *Если для любого* $N = N_k$ *ряд*

$$p_k(g) = \sum_{r,\,i,\,j,\,\alpha} n_r a^{(r)}_{i\alpha} f^N_{r\alpha j} \overline{m^{(r)}_{ij}}(g)$$

равномерно сходится к неотрицательной функции и $A_0 = I$, *то* $\{A_r\}$ *есть преобразование Фурье распределения* $P \in \mathcal{F}(G)$. *Распределение* P *является слабым пределом абсолютно непрерывных распределений с плотностями* $p_k(g)$.

Доказательство. Так как $|m^{(r)}_{ij}(g)| \leqslant 1$, $|a^{(r)}_{ij}| \leqslant \leqslant 1$, то ряд сходится равномерно, и матрицы

$$\Big\{\sum_\alpha a^{(r)}_{i\alpha} f^N_{r\alpha j};\ i, j = 1, 2, \ldots, n_r\Big\} = A^r F^N_r,$$

где

$$F^N_r = \{f^N_{rij};\ i, j = 1, 2, \ldots, n_r\},$$

являются преобразованиями Фурье абсолютно непрерывных распределений с плотностями $p_k(g)$. Но для любого r имеем $F^N_r \to I$ при $N \downarrow e$, так что $A^r F^N_r \to A_r$, чем доказательство завершается.

Можно сформулировать ряд других критериев того, что заданное множество матриц $A^{(r)}$ является преобразованием Фурье распределения вероятностей на $\hat{G}$ (см. замечания 3.2.$_1$). Однако, как и в случае действительной прямой, эти критерии не очень удобно применять в конкретных ситуациях.

Изучение поведения P^{n*} для больших значений n сводится с помощью изоморфизма $P \to \hat{P}$ к изучению поведения $(\hat{P})^n$. Высокие степени матрицы тесно связаны с ее наибольшим по абсолютной величине собственным значением λ, и важно найти критерии того, что $|\lambda| < 1$.

Лемма 3.2.1. *Если некоторое $\hat{P}_r$ $(r \neq 0)$ имеет собственное значение λ на единичной окружности, $|\lambda| = 1$, то носитель $s(P)$ должен содержаться в замкнутой собственной подгруппе A группы G или в некотором классе смежности $g_0 A$.*

Доказательство. Предположим, что $\lambda = e^{i\alpha}$ есть собственное значение $\hat{P}_r$, так что существует нетривиальный вектор z, такой, что $\hat{P}_r z = \lambda z$. Тогда

$$\| z \| = \| \hat{P}_r z \| = \| \int_G M^{(r)}(g) P(dg) z \| \leqslant$$
$$\leqslant \int_G \| M^{(r)}(g) z \| P(dg) \leqslant \| z \|,$$

и равенство имеет место тогда и только тогда, когда

$$s(P) \subset A_\alpha = \{g \mid M^{(r)}(g) z = e^{i\alpha} z\}.$$

Множество $A_0 = \{g \mid M^{(r)}(g) z = z\}$ является замкнутой, собственной [заметим, что $M^{(r)}(g)$ неприводимо] подгруппой. То, что $A_\alpha = g_0 A_0$, $g_0 \in A_\alpha$, завершает доказательство.

Можно показать, что если $s(P)$ не содержится ни в какой такой подгруппе или классе смежности, то $\| \hat{P}_r \| < 1$ (см. замечания $3.2._1$).

Если распределение вероятностей не сосредоточено на какой-нибудь такой подгруппе или классе смежности, то тогда $\hat{P}_r^{n*} = (\hat{P}_r)^n \to 0$ при $n \to \infty$ для любого $r \neq 0$. Каково соответствующее предельное распределение?

Лемма 3.2.2. *Однозначно определенное распределение вероятностей P с преобразованием Фурье*

$$\hat{P}_0 = I, \quad \hat{P}_r = 0, \; r = 1, 2, \ldots,$$

есть нормированная мера Хаара на G.

Доказательство. Для того чтобы убедиться в том, что для меры Хаара $\hat{P}_r = 0$ при $r \neq 0$, достаточно заметить, что элементы $\hat{P}_r$ можно рассматривать как скалярные произведения в $L_2(G)$ элементов $m_{ij}^{(r)}(g)$ и функции, тождественно равной единице, т. е.

$$\{\hat{P}_r\}_{ij} = \int_G m_{ij}^{(r)}(g) P(dg) = (m_{ij}^{(r)}, 1).$$

Утверждение следует тогда из соотношений ортогональности для функций $m_{ij}^{(s)}(g)$.

Если $e \in s(P)$, то $s(P)$ не может содержаться ни в каком классе смежности по какой-либо собственной замкнутой подгруппе. Это приводит нас к следующему результату.

Лемма 3.2.3. *Если $s(P)$ содержит единичный элемент e, то P^{n*} сходится к нормированной мере Хаара на G или на одной из ее замкнутых подгрупп.*

Более законченным предельным результатом для компактных стохастических групп является следующий.

Теорема 3.2.4. *Для заданного распределения вероятностей P предел P^{n*} при $n \to \infty$ существует тогда и только тогда, когда $s(P)$ не содержится ни в каком классе смежности ни по какому замкнутому собственному нормальному делителю группы G. Предел P^{n*} есть нормированная мера Хаара на G.*

Доказательство. Как обычно, предполагается, что G при необходимости модифицируется в наименьшую замкнутую подгруппу, содержащую $s(P)$. Доказательство теоремы в одну сторону аналогично уже проведенному. Действительно, если предел не существует, то можно показать (например, представляя $\hat{P}_r$ в канонической жордановой форме, читатель может предпочесть рассуждение, использующее норму $\hat{P}_r$, см. замечания $3.2._1$), что некоторое $\hat{P}_r$ должно иметь собственное значение $\lambda = e^{i\alpha}$, равное по модулю единице, так что $s(P)$ должно было бы содержаться при $\alpha = 0$ в замкнутой собственной подгруппе A_0 группы G, что исключается. Если $\alpha \not\equiv 0 \pmod{2\pi}$, то $s(P) \subset g_0 A_0 = A_0 g_0$. Введем замкнутое множество $B = \{g \mid gA_0 = A_0 g\}$.

Это — замкнутая подгруппа, содержащая g_0A_0 и $s(P)$, так что должно быть $B = G$, т. е. A_0 есть нормальный делитель.

Для доказательства теоремы в обратную сторону рассмотрим случайный элемент $\gamma_m = g_1g_2 \ldots g_m$, где g_ν имеет вид g_0m; m — случайный элемент в замкнутом собственном нормальном делителе N подгруппы G и $g_0 \overline{\in} N$. В силу того, что N — нормальный делитель, можно, переставив множители, записать γ_n в виде $g_0^n k_n$, где случайные элементы k_n принадлежат N. Для $n = 1, 2, 3, \ldots$ мы получим элементы в множествах

$$g_0N, \; g_0^2N, \; g_0^3N, \; \ldots,$$

и интуитивно ясно, что это несовместимо со сходимостью. Полное доказательство проводится следующим образом. Введем замкнутую коммутативную подгруппу $C=\{g_0^n \mid n = 0, \pm 1, \pm 2, \ldots\}$. Множество CN является замкнутой подгруппой группы G (так как N — нормальный делитель), и, поскольку $s(P) \subset CN$, должно иметь место равенство $CN = G$. Введем гомоморфизм $g \to gN$, отображающий подгруппу C на коммутативную и компактную группу G/N. Используем теперь характер χ на G/N. Мы знаем, что существует характер χ (используется свойство полноты), такой, что $\chi = 1$ на N и $\chi = e^{i\alpha} \neq 1$ на g_0N. Продолжим этот характер на G по формуле $k(g) = \chi(gN)$, так что $k(g) = \chi(g_0N) = e^{i\alpha}$ для $g \in g_0N$. Далее имеем преобразование Фурье

$$\int_G k(g)\,P(dg) = \int_{s(P)} k(g)\,P(dg) = e^{i\alpha},$$

так как $s(P) \subset g_0N$. Но мы не можем получить сходимости $e^{i\alpha n}$ при $n \to \infty$, так что P^{n*} не сходится, что противоречит допущению. Доказательство теоремы закончено (см. замечания 3.2.$_1$).

Легко видеть, что предел P^{n*}, если он существует, является нормированной мерой Хаара.

Это делает возможным решение вопросов, относящихся к сходимости на компактных стохастических группах. Интересен случай, когда $s(P)$ содержит единичный элемент, тогда P^{n*} сходится. Если P симметрично, то $s(P^{2*})$ содержит единичный элемент, так что P^{2n*} сходится.

Естественно возникает вопрос, монотонно ли приближается P^{n*} к некоторому $m \in \mathcal{P}(G)$ при возрастании n.

Теорема 3.2.5. *Введем меру отклонения двух распределений P и Q друг от друга:*

$$d(P, Q) = \sup_E |P(E) - Q(E)|.$$

Если m — такое распределение, что $m * P = m$, то

$$d(P^{(n+1)*}, m) \leqslant d(P^{n*}, m).$$

Доказательство. Имеем

$$P^{(n+1)*}(E) = \int_G P^{n*}(Ey^{-1}) P(dy),$$

$$m(E) = \int_G m(Ey^{-1}) P(dy)$$

и, используя идею, восходящую к Маркову, получаем

$$|P^{(n+1)*}(E) - m(E)| \leqslant \int_G |P^{n*}(Ey^{-1}) - m(Ey^{-1})| P(dy) \leqslant d(P^{n*}, m),$$

что доказывает неравенство.

Польза критерия d ограничена. Если, например, P_n сингулярны относительно m, то $d_n \equiv 1$ и критерий ничего нам не дает.

Можно сказать кое-что о распределениях m, удовлетворяющих соотношению $m * P = m$. Применим преобразование Фурье к этому равенству и получим $\hat{m}_r(\hat{P}_r - I) = 0$. Если все матрицы $\hat{P}_r - I$ при $r \neq 0$ невырожденные, то $\hat{m}_r = 0$ при $r \neq 0$, так что m есть мера Хаара на G. С другой стороны, если для некоторого $r \neq 0$ матрица $\hat{P}_r - I$ имеет собственное значение, равное нулю, то матрица $\hat{P}_r$ должна иметь собственное значение, равное единице.

Но мы знаем, что это возможно только в вырожденном случае, когда при распределении P вся масса сосредоточена на некоторой замкнутой собственной подгруппе группы G.

Критерий d имеет особенно простую геометрическую интерпретацию для абсолютно непрерывных распределений. Пусть мы хотим сравнить меру Хаара μ с $Q \in \mathcal{A}(G)$. Положим $Q(dg)/\mu(dg) = q(g)$. Тогда

$$Q(E) - \mu(E) = \int_E [q(g) - 1]\,\mu(dg),$$

так что

$$d(Q, \mu) = \sup_E |Q(E) - \mu(E)| = \int_F [q(g) - 1]\,\mu(dg),$$

где $F = \{g \mid q(g) > 1\}$. Другими словами, отклонение d между Q и μ есть объем, заключенный между поверхностями функции, тождественно равной единице, и $q(g)$.

К той же проблеме относится и следующее. Предположим, что $P(E) \geqslant cm(E)$, где m — нормированная мера Хаара на группе, E — произвольное борелевское множество и c — константа, $0 < c < 1$. Отсюда вытекает, что $s(P) = G$ и что производная Радона — Никодима $m(dg)/P(dg)$ существует и ограничена. Введем меру $Q = P - cm = D + (1 - c)\,m$, где $D = P - m$. Тогда

$$Q^{n*} = D^{n*} + (1 - c)^n m,$$

так как $D * m = P * m - m = 0$. Следовательно, для любого борелевского множества E

$$P^{n*}(E) = D^{n*}(E) + m(E) = Q^{n*}(E) + [1 - (1 - c)^n]\,m(E),$$

так что

$$P^{n*}(E) - m(E) \geqslant -(1 - c)^n m(E) \geqslant -(1 - c)^n.$$

Взяв дополнительное множество, получим

$$P^{n*}(E) - m(E) \leqslant (1 - c)^n.$$

Этим доказана

Теорема 3.2.6. *Если для некоторой константы* c, $0 < c < 1$,

$$P(E) \geqslant cm(E)$$

для любого борелевского множества E, *то*

$$|P^{n*}(E) - m(E)| \leqslant (1-c)^n$$

(см. замечания 3.2.$_3$).

Можно изучать однородные процессы на G с учетом того, что было сказано в разд. 2.3.

Так, заданное множество функций Q_t, принимающих матричные значения, является преобразованием Фурье $\hat{P}_t$ непрерывного однородного процесса тогда и только тогда, когда выполнено следующее условие. На множестве $\mathcal{U}$ конечных линейных комбинаций функций $u_{ij}(g)$ определяется функционал L, являющийся пределом функционалов L_n вида

$$L_n f(g) = \sum_{\nu=1}^{n} c_\nu^{(n)} [f(g_\nu^{(n)}) - f(e)], \quad c_\nu^{(n)} \geqslant 0 ;$$

(предел понимается в том смысле, что $L_n u \to Lu$ для любой $u \in \mathcal{U}$). Функция Q_t должна быть представима в виде

$$Q_t = \exp(tA), \tag{117}$$

где матрица $A = \{a_{ij}\}$ имеет элементы

$$a_{ij} = Lu_{ij}(g).$$

Перейдем теперь к безгранично делимым законам распределения на G. Нас интересует критерий того, что заданное распределение вероятностей может быть вложено в непрерывный однородный процесс. Говорят, что распределение вероятностей $P \in \mathcal{P}(G)$ вложимо в непрерывный случайный процесс, если существует такойнепрерывный случайный процесс Q_t, что $P = Q_1$.

Теорема 3.2.7. *Распределение P вложимо в непрерывный однородный процесс тогда и только тогда, когда* *существуют такие* $P_n \in \mathcal{P}(G)$, $n = 1, 2, \ldots$, *что*

$$P = P_n^{n*}$$

и

$$\| \hat{P}_n - I \| = O\left(\frac{1}{n}\right).$$

Доказательство. Необходимость этого условия проверяется почти непосредственно. Действительно, если $P = Q_1$, то можно разложить P на множители $Q_{1/n}$, т. е. $P = Q_{1/n}^{n*}$. Рассмотрим непрерывную функцию Q_t, принимающую матричные значения. Так как эти матрицы конечномерны, то эта функция непрерывна также в равномерной операторной топологии, и из полугруппового свойства $\hat{Q}_{t+s} = \hat{Q}_t \hat{Q}_s$ вытекает существование ограниченного инфинитезимального порождающего оператора A. Но тогда

$$\lim_{n \to \infty} n(\hat{Q}_{1/n} - I) = A,$$

так что

$$\| \hat{Q}_{1/n} - I \| \sim \frac{1}{n} \| A \| = O\left(\frac{1}{n}\right),$$

как и утверждалось.

Обратно, предположим, что P может быть разложено на множители P_n, малые в том смысле, что $\| \hat{P}_n - I \| = O(1/n)$. Мы хотим определить Q_t для $t \geqslant 0$. Пусть $r_1, r_2, \ldots$ — неотрицательные рациональные числа. Рассмотрим вероятностные меры $P_n^{[nr_\nu]*}$. Можно найти такую подпоследовательность $n_1, n_2, \ldots$, что слабые пределы

$$\lim_{k \to \infty} P_{n_k}^{[n_k r_\nu]*} = Q_{r_\nu}$$

существуют при $\nu = 1, 2, \ldots$; надо только использовать компактность $\mathcal{P}(G)$ и процедуру диагонального выбора. Заметим, что $Q_1 = P$. Если r и s — неотрицательные рациональные числа, то

$$P_{n_k}^{[n_k(r+s)]*} = P_{n_k}^{[n_k r]*} * P_{n_k}^{[n_k s]*} * P_{n_k}^{\varepsilon *},$$

где ε может принимать только значения 0 и 1. Переходя к пределу, получаем

$$Q_{r+s} = Q_r * Q_s.$$

Чтобы распространить это определение на иррациональные значения параметра, достаточно показать, что $Q_r \to \delta_e$ при $r \to 0$ по рациональным неотрицательным значениям. Но мы имеем

$$\| \hat{Q}_r - I \| = \lim_{k \to \infty} \| \hat{P}_{n_k}^{[n_k r]} - I \|$$

и

$$\| \hat{P}_{n_k}^{[n_k r]} - I \| \leqslant \| \hat{P}_{n_k} - I \| \sum_{0}^{[n_k r]-1} \| \hat{P}_{n_k}^{\nu} \| \leqslant r n_k \| \hat{P}_{n_k} - I \| O(r),$$

что и завершает доказательство.

До сих пор не ясно, можно ли заменить $O(1/n)$ на $o(1)$. Последнее было бы эквивалентно тому, что последовательность P_n слабо сходится к δ_e, а на действительной прямой этого, как известно, достаточно.

Можно показать также достаточность для симметричных распределений P_n, надо только исследовать собственные значения эрмитовых матриц $\hat{P}_n$. В общем компактном случае это рассуждение не проходит; неприятность проистекает из того, что корень n-й степени является многозначной функцией на комплексной плоскости. Иными словами, трудность заключается в разделении ветвей логарифмической функции.

Можно упомянуть, что если распределение P безгранично делимо, то $\hat{P}$ — невырожденная матрица.

Замечание. Если потребовать больше, а именно, чтобы ограничение $O(1/n)$ было выполнено и для норм $\| P_n - \delta_e \|$, то Q_t становится сложным пуассоновским процессом.

Можно выразить спектр оператора

$$Tf(g) = \int_G f(gh) P(dh)$$

через спектры матриц $\hat{P}_r$. Сделаем это для симметричных распределений вероятностей.

Теорема 3.2.8. *Спектр оператора T, соответствующего симметричному распределению вероятностей P, дискретен. Он состоит из собственных значений всех матриц $\hat{P}_r$, причем собственное значение матрицы $\hat{P}_r$ должно считаться d_r раз, если $\hat{P}_r$ есть квадратная матрица порядка d_r.*

Доказательство. Пусть λ есть собственное значение матрицы $\hat{P}_r$. Так как матрица $\hat{P}_r$ эрмитова, то она может быть приведена к диагональному виду, и значение λ будет стоять на диагонали. Другими словами, мы выбираем новую подходящую прямоугольную систему координат. Обозначим функции $m_{ij}^{(r)}(g)$, выраженные в новых координатах, через $n_{ij}^{(r)}(g)$. Конечно, новые функции обладают свойствами, аналогичными тем, которым удовлетворяли старые, такими, как ортогональность и т. п. Имеем

$$\left\{\int\limits_G n_{ij}^{(r)}(g)\,P(dg);\ i,j=1,\,2,\,\ldots,\,d_r\right\}=\begin{pmatrix} a_1 & & & 0 \\ & a_2 & & \\ & & \ddots & \\ 0 & & & a_{d_r}\end{pmatrix},$$

где, скажем, $a_1=\lambda$. Тогда

$$T_{n_{i1}}^{(r)}(g)=\int\limits_G n_{i1}^{(r)}(gh)\,P(dh)=$$

$$=\sum_{j=1}^{d_r} n_{ij}^{(r)}(g)\int\limits_G n_{j1}^{(r)}(h)\,P(dh)=a_1 n_{i1}^{(r)}(g).$$

Отсюда вытекает, что $n_{i1}^{(r)}(g)$ есть собственная функция оператора T для $i=1,\,2,\,\ldots,\,d_r$. Но тогда λ — собственное значение оператора T и должно считаться d_r раз. Поскольку функции $n_{ij}^{(r)}(g)$, как и $m_{ij}^{(r)}(g)$, образуют полную систему, то этими собственными значениями исчерпывается спектр оператора T.

3.3. Коммутативные локально компактные стохастические группы

Для случая групп, которые коммутативны, но не обязательно компактны, а только локально компактны, неприводимые унитарные представления имеют следующий вид.

Они не обязаны больше образовывать счетное множество, но это усложнение уравновешивается тем фактом, что они одномерны, $n_\tau \equiv 1$. *Характер* $\gamma = (g, \gamma)$ есть непрерывная комплекснозначная функция на G, удовлетворяющая равенствам

$$(g+h,\ \gamma) = (g,\ \gamma)\cdot(h,\ \gamma),$$
$$|(g,\ \gamma)| \equiv 1.$$

Бинарная групповая операция записывается в этом случае как сложение. Введем *двойственную группу* Γ, состоящую из всех характеров, с окрестностями нуля вида

$$\{\gamma \,|\, |1-(g,\ \gamma)| < \varepsilon\}$$

для всех g из компактного множества C и для положительного ε, и определим сложение соотношением

$$(g,\ \gamma+\delta) = (g,\ \gamma)\cdot(g,\ \delta).$$

Известно, что Γ локально компактна и коммутативна и имеет двойственной группой группу G. Дискретной группе G соответствует компактная двойственная группа, а компактная группа G имеет дискретную двойственную группу. Для любого $g \neq 0$ существует такое $\gamma \in \Gamma$, что $(g, \gamma) \neq 1$.

Для любого $P \in \mathcal{P}(G)$ определим преобразование Фурье распределения P как комплекснозначную функцию на Γ:

$$\hat{P} = \hat{P}(\gamma) = \int_G (g,\ \gamma)\, P(dg), \quad \gamma \in \Gamma.$$

Теорема 3.3.1. *На коммутативной локально компактной группе преобразование Фурье P обладает следующими свойствами.*

а) *P однозначно определено преобразованием $\hat{P}$.*

б′) *Пусть последовательность распределений вероятностей $P_1, P_2, \ldots$ сходится к $P \in \mathcal{P}(G)$. Тогда $\hat{P}_n \to \hat{P}$.*

б″) *Пусть последовательность распределений вероятностей $P_1, P_2, \ldots$ имеет преобразования Фурье $\hat{P}_n(\gamma)$, сходящиеся к непрерывной функции $\hat{P}(\gamma)$, $\gamma \in \Gamma$. Тогда $\hat{P}$ есть преобразование Фурье некоторого распределения P, и имеет место слабая сходимость $P_n \to P$.*

в) $|\hat{P}(\gamma)| \leqslant 1$.

г) $\widehat{P_1 * P_2} = \hat{P}_1 \cdot \hat{P}_2$.

д) $\hat{P}(\gamma)$ *непрерывна.*

Доказательство. Доказательство утверждений а), б′), в) и г) аналогично доказательству соответствующих утверждений в разд. 3.2. Будем основывать доказательство б″) на теореме 3.3.2, утверждающей, что функции $\hat{P}_n(\gamma)$ являются положительно определенными. Предельная непрерывная функция $\hat{P}(\gamma)$ тогда тоже является положительно определенной и должна быть преобразованием Фурье некоторого распределения $P \in \mathcal{P}(G)$. Известно (см. замечания 5.1), что можно аппроксимировать константу 1 равномерно на любом компактном множестве C непрерывными положительно определенными функциями p, равными нулю вне компактных множеств, а в 0 принимающими значение 1. Пусть теперь P_{n_ν} есть последовательность, L-слабо сходящаяся к мере Q, $Q(G) \leqslant 1$. Нужно показать, что $Q(G) = 1$. Так как функция p является положительно определенной, то она может быть записана в виде[1]

$$p(g) = \int_\Gamma (g, \gamma)\, \mu(d\gamma), \text{ где } \mu \in \mathcal{P}(\Gamma).$$

Следовательно,

$$\int_\Gamma p(g) P_{n_\nu}(dg) = \int_G \int_\Gamma (g, \gamma)\, \mu(d\gamma) P_{n_\nu}(dg) =$$

$$= \int_\Gamma \hat{P}_{n_\nu}(\gamma)\, \mu(d\gamma) \to \int_\Gamma \hat{P}(\gamma)\, \mu(d\gamma) = \int_G p(g) P(dg),$$

и мы имеем

$$\int_G p(g) P_{n_\nu}(dg) \to \int_G p(g) Q(dg).$$

Возьмем теперь C очень большим и получим $Q(G) = P(G) = 1$. Оставшаяся часть доказательства б″) стандартна.

---

[1]) См. теорему 3.3.2. ниже.— *Прим. ред.*

Легко доказать утверждение д). Действительно, для любого $\varepsilon > 0$ можно выбрать такое компактное множество C, что $P\ (\bar{C}) < \varepsilon$ и

$$|\hat{P}(\gamma)-\hat{P}(\gamma')| \leqslant 2\varepsilon + \int_C |(g, \gamma)-(g, \gamma')|\, P(dg),$$

где правая часть может быть сделана малой, если γ' выбирается в малой окрестности γ.

Важное соотношение между положительно определенными функциями и преобразованиями Фурье распределений вероятностей дается следующей знаменитой теоремой, которую мы приводим без доказательства (см. замечания 3.3).

Теорема 3.3.2. *Если $p\ (\gamma)$ — непрерывная положительно определенная функция, такая, что $p\ (0) = 1$, то она является преобразованием Фурье $\hat{\mu}$ некоторой меры $\mu \in \mathcal{P}(G)$,*

$$p(\gamma) = \int_\Gamma (g, \gamma)\, \mu\,(dg),$$

и обратно.

Теперь мы можем перейти к изучению идемпотентных распределений вероятностей. Предположим, что распределение $P \in \mathcal{P}\ (G)$ идемпотентно, $P^{2*} = P$, и будем, как обычно, предполагать, что G выбирается как наименьшая замкнутая подгруппа, содержащая $s\ (P)$. Если бы преобразование Фурье $\hat{P}\ (\gamma)$ было постоянным на классах смежности некоторой нетривиальной замкнутой подгруппы $\Gamma_0 \subset \Gamma$, т. е.

$$\int_G (g, \gamma+\delta)\, P\,(dg) = \int_G (g, \gamma)\cdot(g, \delta)\, P\,(dg) =$$

$$= \int_G (g, \gamma)\, P\,(dg), \quad (\gamma \in \Gamma,\ \delta \in \Gamma_0),$$

то тогда бы $(g, \delta)\ P\ (dg) = P\ (dg)$ при $\delta \in \Gamma_0$. Следовательно, $(g, \delta) = 1$ для g из $s\ (P)$, так что этот носитель должен содержаться в аннулирующей подгруппе $\{g \mid (g, \delta) = 1$ при $\delta \in \Gamma_0\}$, соответствующей Γ_0, а это исключается нашими предположениями. Определяя теперь комплексную меру m_α, $\alpha \in \Gamma$, соотношением $m_\alpha(dg) = (g, \alpha) P(dg)$, получаем, что $\hat{m}_\alpha(\gamma) = \hat{P}(\gamma + \alpha) \not\equiv \hat{P}(\gamma)$, так что $m_\alpha \neq P$, если $\alpha \neq 0$. Выберем теперь столь большое компактное подмножество C группы G, чтобы $P(\overline{C}) < 1/4$. Рассмотрим открытое подмножество V группы Γ:

$$V = \left\{\gamma \,\|\, 1 - (g, \gamma) | < \frac{1}{3}, \; g \in C\right\}.$$

Имеем для $\alpha \in V$

$$\|P - m_\alpha\| = \int_G |1 - (g, \alpha)| P(dg) \leqslant \frac{1}{3} P(C) + 2P(\overline{C}) < \frac{5}{6} < 1.$$

С другой стороны, $\|P - m_\alpha\| \geqslant \sup_\gamma |\hat{P}(\gamma) - \hat{m}_\alpha(\gamma)| = \sup_\gamma |\hat{P}(\gamma) - \hat{P}(\gamma + \alpha)| \geqslant 1$ при $\alpha \neq 0$, так как $\hat{P}$ принимает только значения 0 и 1, поскольку $(\hat{P})^2 = \hat{P}$. Это показывает, что V содержит только один элемент 0, так что Γ дискретна и, следовательно, G компактна. Это доказывает следующую теорему (см. замечания 3.3.$_1$), которая сводит изучение идемпотентных мер к компактному случаю.

Теорема 3.3.3. *Идемпотентная вероятностная мера на коммутативной локально компактной группе необходимо сосредоточена на некоторой компактной подгруппе.*

Нетрудно модифицировать лемму 3.2.1 для нашего случая. Действительно, если $\hat{P}(\gamma) = e^{i\alpha}$, то вся масса при распределении P должна находиться на множестве $A_\alpha = \{g \mid (g, \gamma) = e^{i\alpha}\}$, которое является классом смежности замкнутой подгруппы A_0.

Лемма 3.3.1. *Вероятностная мера P на локально компактной коммутативной группе, преобразование Фурье $\hat{P}(\gamma)$ которой принимает значение $e^{i\alpha}$, по модулю равное единице, имеет массу, сосредоточенную на некотором классе смежности A_α замкнутой подгруппы A_0 группы G. При этом $\hat{P}(\gamma) = e^{i\alpha}$ для всех γ из некоторого класса смежности замкнутой подгруппы B группы Γ, аннулирующей A_α, $(g, \gamma) = 1$ для всех $g \in A_\alpha$, $\gamma \in B$.*

На локально компактной (но не компактной) коммутативной группе мера Хаара бесконечна и не может быть нормированием превращена в вероятностную меру. Поэтому теорему 3.2.5 непосредственно нельзя применить. В некотором смысле остается справедливым, что *композиция стремится выравнять распределения* на G. Действительно, рассмотрим класс $\mathcal{A}(G)$ абсолютно непрерывных распределений и определим *меру концентрации* соотношением

$$C(P)=\int\limits_G p^2(g)\,\mu(dg),$$

где $p(g)=P(dg)/\mu(dg)$. Рассмотрим распределение P_1 конечной концентрации $C(P_1)$ и подвергнем его композиции с P_2. Концентрация $P_1 * P_2$ равна тогда

$$C(P_1*P_2)=\int\limits_G\left[\frac{P_1*P_2(dg)}{\mu(dg)}\right]^2\mu(dg)=\int\limits_\Gamma|\hat P_1(\gamma)\hat P_2(\gamma)|^2 m(d\gamma)\leqslant$$

$$\leqslant\int\limits_\Gamma|\hat P_1(\gamma)|^2 m(d\gamma)=\int\limits_G\left[\frac{P_1(dg)}{\mu(dg)}\right]^2\mu(dg)=C(P_1)$$

в силу соотношения Парсеваля; $m(d\gamma)$ обозначает соответствующим образом нормированную меру Хаара на группе характеров Γ.

Теорема 3.3.4. *При принятом выше определении композиция не увеличивает концентрации* (см. замечания $3.3._2$).

Как мы видели, анализ Фурье распределений вероятностей на локально компактной коммутативной группе вполне аналогичен тому, который используется на действительной прямой. Примечательно, что наши знания о вероятностных предельных теоремах на таких группах не отличаются полнотой (см. замечания $3.3._4$).

В этой связи можно было бы поставить вопрос о мере рассеяния, ведущей себя более или менее аналогично дисперсии. Непосредственное обобщение понятия дисперсии возможно в линейном пространстве (см. 6.1), но мы можем прийти к интересным понятиям и в отсутствие линейной структуры.

Пусть сначала G — локально компактная коммутативная группа. Для заданного распределения вероятностей

P на G образуем величину

$$d_\gamma(P) = -\log|\hat{P}(\gamma)|,$$

которая либо является неотрицательным действительным числом, либо принимает значение $+\infty$. Для любого характера $\gamma \in \Gamma$ функционал $d_\gamma(P)$, определенный на $\mathcal{P}(G)$, является аддитивным:

$$d_\gamma(P_1 * P_2) = -\log|\widehat{P_1 * P_2}| = -\log|\hat{P}_1| - \log|\hat{P}_2| = \\ = d_\gamma(P_1) + d_\gamma(P_2).$$

Если мы применим к этим d_γ (где γ изменяется в двойственной группе Γ) некоторую линейную операцию, то придем к некоторому аддитивному функционалу $d(P)$, который примем в качестве нашей меры рассеяния $d(P) = \mathcal{L} d_\gamma(P)$.

Мы пока не уточняли, как должен пониматься линейный оператор $\mathcal{L}$. Это можно сделать с помощью борелевской меры H, определенной на Γ:

$$d(P) = \mathcal{L}\, d_\gamma(P) = \int_\Gamma d_\gamma(P)\, H(d_\gamma).$$

Удобно предположить, что $s(H) = \Gamma$. Тогда имеет место

Теорема 3.3.5. *Функционал*

$$d(P) = -\int_\Gamma \log|\hat{P}(\gamma)|\, H(d\gamma)$$

является мерой рассеяния в том смысле, что

а) $0 \leqslant d(P) \leqslant \infty$,

б) $d(P_1 * P_2) = d(P_1) + d(P_2)$,

в) $d(P) = 0$ *тогда и только тогда, когда* P *сводится к вырожденному распределению* δ_g,

г) *при композиции* $d(P)$ *не убывает.*

Доказательство теоремы совсем просто. Утверждения а), б) и г) проверяются непосредственно. Для того чтобы убедиться в выполнении утверждения в), заметим, что из $d(P) = 0$ вытекает, что

$$|d_\gamma(P)| = 0$$

для всех $\gamma \in \Gamma$, так как носителем меры H является вся группа Γ, а $d_\gamma(P)$ — непрерывная функция от γ. Образуем симметризованную вероятностную меру $S = P * \overline{P}$, соответствующую случайному элементу gh^{-1}, где g и h независимы и имеют распределение P. Преобразование Фурье меры S имеет вид $\hat{P}(\hat{\overline{P}}) = |\hat{P}|^2 = 1$, так что S — вырожденное распределение δ_e. Следовательно, распределение P должно быть сосредоточено на некотором постоянном элементе g, что доказывает утверждение. Обратное очевидно.

Может оказаться интересиым рассмотрение других форм оператора $\mathscr{L}$, особенно таких, которые локальны в единичном элемеите e. Под этим мы подразумеваем, что оператор использует только те $d_\gamma(P)$, у которых γ близко к e. Далее это не развивается, но читатель заметит, что это связано со способом введения обычной дисперсии. Другой проблемой, оставленной без рассмотрения, являются свойства «сходимости», определяемой соотношением $d(P_n) \to 0$.

Если G есть компактная группа, то сказанное выше должно быть модифицировано, так как преобразование Фурье распределения P в этом случае состоит из последовательности матриц $\hat{P}_0, \hat{P}_1, \ldots$. Если заменить в определеиии $d_\gamma(P)$ абсолютную величину нормой, то мы приходим к мере рассеяния вида

$$d(P) = -\sum_{n=0}^{\infty} H_n \log \|\hat{P}_n\|,$$

где H_ν — положительные константы. Этот функционал уже не обязательно аддитивен, но так как

$$\|\widehat{P_n^{(1)} * P_n^{(2)}}\| = \|\hat{P}_n^{(1)} \hat{P}_n^{(2)}\| \leqslant \|\hat{P}_n^{(1)}\| \cdot \|\hat{P}_n^{(2)}\|,$$

то

$$d(P_1 * P_2) \geqslant d(P_1) + d(P_2).$$

Заметим, что нормированная мера Хаара, как и следует ожидать, имеет бесконечное рассеяние.

Наконец, если G есть общая локально компактная группа, то нужно исходить из преобразований Фурье типа, используемого в гл. 5. Меры рассеяния при этом изучаться не будут.

В случае, когда G — коммутативная группа с инвариантной метрикой $(g', g'') = (g' + h,\ g'' + h)$, можно ввести меру рассеяния следующим образом (если G — группа Ли, то можно представлять себе (g', g'') как инвариантную риманову метрику). В аддитивных обозначениях можно записать (g', g'') в виде $\| g' - g'' \|$. Заметим, что эта норма не является линейной, так как для постоянного скаляра c не обязательно имеет смысл cg. Определим

$$d(P) = \inf_{g \in G} \left[\int_G \| g - h \|^2 P(dh) \right]^{1/2}.$$

Если допустить значение $+\infty$, то эта величина всегда определена. Мы не можем утверждать, вообще говоря, что существует единственное значение g, при котором реализуется минимум (представьте себе компактную группу с нормированной мерой Хаара). Имеем $0 \leqslant d(P) \leqslant +\infty$, причем $d(P) = 0$ тогда и только тогда, когда P — вырожденное распределение, сосредоточенное на единственном элементе группы G. Далее,

$$d(P' * P'') \leqslant d(P') + d(P'').$$

Для того чтобы убедиться в этом, учтем, что при $g = g' + g''$

$$\| h' + h'' - g \| \leqslant \| h' - g' \| + \| h'' - g'' \|.$$

Здесь мы рассматриваем h' и h'' как два независимых элемента стохастической группы с распределениями вероятностей соответственно P' и P''. Возводя в квадрат и интегрируя, получаем нужное соотношение с помощью неравенства Шварца. Заметим, что равенство в общем случае не обязано иметь места. Типичный пример: если G — компактная группа, а $P' = P''$ — нормированная мера Хаара, то

$$d(P' * P'') = d(P') < 2d(P')$$

(см. также замечания $3.3._3$).

Так же, как в разд. 3.2, может быть сформулирована и доказана

Теорема 3.3.6. *Пусть P — симметричное распределение вероятностей на локально компактной коммутатив-*

ной группе G. Вероятностный оператор

$$Tf(g)=\int\limits_G f(gh)\,P(dh)$$

имеет своим спектром множество значений преобразования Фурье $\hat{P}(\gamma)$.

Доказательство предоставляется читателю.

3.4. Примеры

Рассмотрим циклическую группу G из m элементов $0, 1, 2, \ldots, m-1$ с групповой операцией сложения по модулю m. Характеры $(g, \gamma)=\exp\left(2\pi i\,\frac{g\gamma}{m}\right)$, $\gamma = 0, 1, 2, \ldots, m-1$, образуют группу Γ. Распределение P вероятностей $P(g=\nu)=p_\nu$, $\nu=0, 1, 2, \ldots, m-1$, на G имеет преобразование Фурье

$$\hat{P}(\gamma)=\sum_{\nu=0}^{m=1} p_\nu \exp\left[\frac{2\pi i\nu\gamma}{m}\right].$$

Последовательность композиций P^{n*} с преобразованиями Фурье $[\hat{P}(\gamma)]^n$ сходится, очевидно, к равномерному распределению с преобразованием Фурье, равным единице для $\gamma=0$ и нулю в остальных случаях, тогда (и только тогда), когда $|\hat{P}(\gamma)|<1$ при $\gamma\neq 0$. Следовательно, эта сходимость имеет место, если носитель P не содержится в классе смежности какой-либо подгруппы группы G, а любая подгруппа группы G является, конечно, циклической группой, порядок которой есть делитель m.

Простые особенности анализа Фурье на $\mathcal{P}(G)$ делают возможным более детальное исследование этой стохастической группы.

Группа вращений окружности T^1 почти столь же проста. Пусть $G=T^1=\{g\,|\,0\leqslant g<2\pi\}$ с обычным сложением и идентификацией по модулю 2π. Характерами являются $e^{i\gamma g}$, $\gamma=\ldots -1, 0, 1, \ldots,$ так что преобразование

Фурье имеет вид

$$\hat{P}(\gamma)=\int_0^{2\pi} e^{i\gamma g} P(dg).$$

Последовательность P^{n*} сходится, как и раньше, к равномерному распределению, если P не содержится в классе смежности какой-нибудь подгруппы. Это полностью решает вопрос сходимости, так как подгруппы являются конечными циклическими группами.

Мы предоставляем читателю возможность рассмотреть эти два случая в свете того, что говорилось в разд. 3.2 и 3.3. Изучение родственных групп T^r не вызывает никаких дополнительных затруднений.

Для конечных групп безгранично делимые распределения могут быть охарактеризованы следующим образом (см. замечания 3.4). Пусть G есть конечная группа порядка n и H — произвольная подгруппа группы G. Обозначим через μ_H нормированное равномерное распределение на H. Рассмотрим групповую алгебру A, в которой $f\cdot h$ обозначает композицию функций f и h, определенных на G.

Теорема 3.4.1. *Выберем такой элемент $f(g)\in A$, что*

1) *$f(g)$ является H-инвариантным:*

$$fh=hf=f, \quad h\in H;$$

2) $\sum\limits_g f(g)=0;$

3) *$f(g)\geqslant 0$ при $g\,\overline{\in}\,H$,*

и определим

$$P=\exp_H f=\mu_H+\sum_1^{\infty}\frac{f^\nu}{\nu!}\,.$$

Эти и только эти P могут быть разложены: $P=P_n^{n*}$, $n=1,\ 2,\ \ldots$.

Если добавить условие $P_n\to\delta_e$, то получим представление

$$P=\exp f$$

в виде сложного пуассоновского распределения; здесь $H=\{e\}$.

Глава 4

СТОХАСТИЧЕСКИЕ ГРУППЫ ЛИ

4.1. Предварительные сведения о группах Ли

В дифференцируемом многообразии можно рассчитывать на построение исчисления вероятностей и описание распределений вероятностей в терминах производных, дифференциальных уравнений и т. п. (см. замечания 4.1.$_1$). Однако в соответствии с основной темой этой книги мы ограничимся группами Ли. Чтобы не затемнять основных идей усложненными, хотя по существу и ясными, рассуждениями, связанными с предельными переходами, мы рассмотрим основные шаги доказательств и только кратко набросаем остающиеся детали. Дальнейшие подробности читатель может найти в работах, цитированных в замечаниях.

Прежде всего введем необходимую терминологию и обозначения. Пусть G есть сепарабельная группа Ли размерности d. Обозначим через $\Lambda(G)$ соответствующую *алгебру Ли инфинитезимальных правых переносов* группы G. Введем множество C, состоящее из непрерывных функций на одноточечном компактном расширении G_c группы G; это банахово пространство с нормой $\|f\| = \max\limits_g |f(g)|$. Для $f \in C$ и инфинитезимального преобразования $Y \in \Lambda$ определим

$$Yf = \lim_{h \downarrow 0} \frac{R_{\eta_{Y(h)}}f - f}{h},$$

где

$$\eta_Y(t) = \exp(tY)$$

и где предел должен существовать в топологии C, т. е. равномерно на G_c. Обозначим через C_1 множество $f \in C$, для которых Yf определено. Пусть C_k есть множество таких f, что $Y_1Y_2 \ldots Y_kf$ определено для любых Y_1, $Y_2, \ldots, Y_k \in \Lambda(G)$. Сделаем его банаховым пространством, введя норму

$$\|f\|_k = \|f\| + \sum_i \|X_i f\| + \ldots + \sum_{i_1, \ldots, i_k} \|X_{i_1} X_{i_2} \ldots X_{i_k} f_k\|,$$

где $X_1, X_2, \ldots, X_k$ — выбранный базис в $\Lambda(G)$. Точно так же введем аналогичные пространства C'_k, исходя из левых (а не правых) переносов. Все эти пространства плотны в C. Мы используем в окрестности e канонические координаты $x_1, x_2, \ldots, x_d$, связанные с $X_1, X_2, \ldots, X_d$, и предполагаем также, что x_i распространяются на G_c с $x_i \in C_2$. В окрестности e введем функцию

$$\varphi(g) = \sum_1^d x_i^2(g).$$

Можно распространить φ на G_c так, чтобы φ принадлежала C_2 и была ограничена снизу положительным числом вне каждой окрестности e.

Введем функционалы

$$D_i f = \lim_{h \downarrow 0} \frac{f(\eta_{X_i}(h)) - f(e)}{h},$$

$$D_{ij} f = \lim_{h \downarrow 0} \frac{f(\eta_{X_i}(h)\,\eta_{X_j}(h)) - f(\eta_{X_i}(h)) - f(\eta_{X_j}(h)) + f(e)}{h^2},$$

если эти пределы существуют; в C_2 они дают соответственно значения $X_i f(e)$ и $X_i X_j f(e)$.

Объектом нашего исследования будут являться борелевские вероятностные меры P на G_c и полугруппы P_t таких вероятностных мер, $P_t * P_s = P_{s+t}$; $s, t \geqslant 0$. P_t должны быть непрерывны в смысле слабой сходимости вероятностных мер на G_c (заметим, что это не то же самое, что слабая сходимость на G), $P_h \to \delta_e$ при $h \downarrow 0$.

Соответствующие операторы

$$T_t f(g) = \int_{G_c} f(gh)\, P_t(dh)$$

определены для $f \in C$ и образуют сильно непрерывную полугруппу.

4.2. Однородные процессы на группах Ли

Если только инфинитезимальный порождающий оператор M полугруппы T_t не является ограниченным, то мы сталкиваемся с обычными трудностями в силу того, что M не определен всюду на C. Но теперь в нашем распоряжении более подробные сведения о группе, и мы можем говорить о дифференцируемых функциях и т. п. Это делает возможным описание M в терминах дифференциальных и интегральных операторов. Заинтересованный читатель может сравнить сказанное ниже с обычным выводом формулы Леви — Хинчина на действительной прямой.

Покажем сначала, что M может быть выражен некоторым образом в подмножестве множества C. Далее нужно показать, что M, заданный на этом подмножестве, действительно определяет T_t. Наконец, мы покажем, что любой M этого вида порождает непрерывный однородный процесс.

Рассмотрим однородный случайный процесс (см. разд. 2.2), принимающий значения из G_c, с распределениями вероятностей P_t и вероятностными операторами T_t. Так как T_t образуют сильно непрерывную полугруппу, инфинитезимальный оператор N полугруппы T_t определен на плотном подмножестве $\mathcal{D}$ множества C. Одако T_t также сильно непрерывна в C_2' (см. замечания 4.2.$_2$) и удобно ввести N', положив

$$N'f = \lim_{h \downarrow 0} \frac{T_h f - f}{h},$$

на некотором подмножестве $\mathcal{D}'$, плотном в C_2'.

Немного изменим теперь функцию $\varphi(g)$. Для любого $\varepsilon > 0$ можно найти такую $\varphi' \in \mathcal{D} \cap C_2'$, что $\| \varphi - \varphi' \|_2' < \varepsilon$, $\varphi'(e) = D_i\varphi' = D_{ij}\varphi' = 0$ при $i \neq j$ и $D_{ii}\varphi' = 2$, и φ' положительна на $G_c - e$.

Докажем теперь, что предел

$$Af = \lim_{h \downarrow 0} \frac{T_h f(e) - f(e)}{h}$$

существует для каждой $f(g)$ из класса $\mathscr{E}$ функций из C, являющихся дважды непрерывно дифференцируемыми. Для заданного положительного δ, используя обычное рассуждение, выбираем такую $f_\delta \in \mathscr{D} \cap C_2$, что $|f(g) - f_\delta(g)| \leqslant \delta\varphi'(g)$. Но тогда

$$\frac{1}{h}\int\limits_{G_c}[f(g)-f(e)]\,P_h(dg)=\frac{1}{h}\int\limits_{G_c}[f(g)-f_\delta(g)]\,P_h(dg)+\\+\frac{1}{h}\int\limits_{G_c}[f_\delta(g)-f_\delta(e)]\,P_h(dg),$$

так как $f_\delta(e) = f(e)$. Предел второго слагаемого существует, так как $f_\delta \in \mathscr{D} \cap C_2$. Но

$$\frac{1}{h}\int\limits_{G_c}[f(g)-f_\delta(g)]\,P_h(dg)\leqslant\frac{\delta}{h}\int\limits_{G_c}\varphi'(g)\,P_h(dg),$$

и так как $\varphi' \in \mathscr{D}$, предел $\frac{1}{h}\int\limits_{G_c}\varphi'(g)\,P_h(dg)$ существует. Так как δ может быть взято сколь угодно малым, то отсюда следует наше утверждение. Можно применить A к функциям $x_i(g)$ и $x_i(g)\,x_j(g)$. Обозначим $b_i = Ax_i$ и $b_{ij} = Ax_ix_j$. Функционал A так важен потому, что, как мы увидим, инфинитезимальный порождающий оператор $\{T_t\}$ может быть выражен через A.

Введем теперь меры, зависящие от параметра t:

$$\eta_t(E)=\frac{1}{t}\int\limits_{E}\varphi(g)\,P_t(dg).$$

Они равномерно ограничены (это следует из того, что $\lim\limits_{t\downarrow 0}\frac{1}{t}\int\limits_{G_c}\varphi(g)\,P_t(dg)$ существует) и при $f \in \mathscr{E}$ интегралы

$$\int\limits_{G_c}f(g)\,\eta_t(dg)=\frac{1}{t}\int\limits_{G_c}f(g)\,\varphi(g)\,P_t(dg)$$

сходятся при $t \to 0$. Таким образом, η_t сходятся к некоторой мере η на G_c при $t \to 0$, так что

$$\lim_{h\downarrow 0}\frac{1}{t}\int\limits_{G_c}f(g)\,\varphi(g)\,P_t(dg)=\int\limits_{G_c}f(g)\,\eta(dg)$$

для всех $f \in C$.

Теперь мы можем получить явное представление A. Если $f \in \mathscr{E}$ и если обозначить $c = f(e)$, $c_i = D_i f$ и $c_{ij} = D_{ij} f$, то функция

$$h(g) = \frac{1}{\varphi(g)} [f(g) - c - \sum c_i x_i - \sum c_{ij} x_i x_j]$$

принадлежит C и равна нулю при $g = e$. Переписываем определение A:

$$Af = A(c + \sum c_i x_i + \sum c_{ij} x_i x_j) + A\varphi h = \sum c_i b_i + \\ + \sum c_{ij} b_{ij} + A\varphi h = \sum c_i b_i + \sum c_{ij} b_{ij} + \\ + \int\limits_{G_c - e} \frac{1}{\varphi(g)} [f(g) - c - \sum c_i x_i(g) - \sum c_{ij} x_i(g) x_j(g)] \, \eta(dg).$$

Интегралы $\int\limits_{G_c - e} \frac{x_i x_j}{\varphi} \eta(dg)$ существуют, так как подынтегральные выражения ограничены, и мы можем опустить эти члены, если согласиться изменить определение b_{ij}. Следовательно, мы можем написать

$$Af = \sum a_i D_i f + \sum a_{ij} D_{ij} f + \\ + \int\limits_{G_c - e} \frac{1}{\varphi(g)} [f(g) - f(e) - \sum D_i f(g) x_i(g)] \, \eta(dg),$$

где коэффициенты a_{ij} могут быть выбраны симметричными, $a_{ij} = a_{ji}$ (это возможно, поскольку $D_{ij} - D_{ji}$ может быть выражено линейно через D_ν). Матрица $\{a_{ij}\}$ должна быть также неотрицательно определенной. Действительно, операторы T_t неотрицательны, и, следовательно, для любой неотрицательной $f \in \mathscr{E}$ с $f(e) = D_i f = 0$ мы должны иметь

$$\sum a_{ij} D_{ij} + \int\limits_{G_c - e} \frac{1}{\varphi} f(g) \, \eta(dg) = \lim_{h \downarrow 0} \frac{1}{h} T_h f \geqslant 0.$$

Выберем теперь такую f, равную $[\sum z_i x_i(g)]^2 \geqslant 0$ в окрестности e, но столь малую, что величина интеграла пренебрежимо мала. Но тогда квадратичная форма $\sum a_{ij} z_i z_j$ должна быть неотрицательной, чтобы избежать противоречия. Пусть теперь A есть описанный выше функционал, но рассматриваемый теперь на C_2. Величины a_i, a_{ij} и η выбираются так, чтобы $Nf(e) = Af$. Если N есть сужение на C_2 инфинитезимального порождающего оператора некоторой неотрицательной полугруппы T'_t, то $T_t = T'_t$, так как оба они приводят к A на C_2. Чтобы проверить последний шаг, мы используем лемму, которая в нескольких, слегка различных, вариантах встречается при изучении параболических дифференциальных уравнений. Пусть $f(t, g)$ есть непрерывная функция на $[0, \infty) \times G_c$, такая, что для любой пары (t, g_0), удовлетворяющей условиям

$$t > 0, \ f(t, g_0) = \min_g f(t, g),$$

производная $\partial f/\partial t$ существует и неотрицательна. Тогда $\min f(t, g) \geqslant \min f(0, g)$ для всех t (см. замечания 4.2.$_1$). Здесь мы хотим доказать, что знание A на C_2 определяет $T_t f$ для всех t. Достаточно рассмотреть f, равные константе около ω и принадлежащие $\mathscr{E}$. Но тогда $T_t f \in \mathscr{E}$, так как $f \in C'_2$ и T_t отображает C'_2 в себя (см. замечания 4.2.$_2$). Теперь мы интерпретируем $f(0, g)$ как $f(g)$, а $f(t, g)$ как $T_t f(g)$. Условия леммы выполнены:

$$\begin{cases} \dfrac{\partial f(t, \omega)}{\partial t} = 0; \ t > 0, \\ \dfrac{\partial f(t, g)}{\partial t} = \lim\limits_{h \downarrow 0} \dfrac{T_h f(t, g) - f(t, g)}{h} = A(L_{g^{-1}} f(t, g)); \ g \in G, \ t > 0. \end{cases}$$

Если теперь $f_1(t, g)$ и $f_2(t, g)$ — функции, обладающие этими свойствами, то их разность $f_3 = f_1 - f_2$ должна удовлетворять указанному выше условию и $f_3(0, g) = 0$, $g \in G$, $f_3(t, \omega) = 0$, $t > 0$. Но если g_0 — точка, в которой $f_3(t, g)$ достигает минимума, то член первого порядка в A обращается в нуль, а сумма членов второго порядка, так же как и слагаемое с интегралом, неотрицательна. Но тогда мы должны иметь $f_3(t, g) \geqslant \min\limits_g f_3(0, g) = 0$. Вместе с f_3 должно быть решением и $-f_3$, так что $-f_3 \geqslant 0$, что приводит нас к желаемому результату. Следовательно, мы знаем вид инфинитезимального порождающего оператора N полугруппы T_t на C_2, причем N на C_2 определяет полугруппу однозначно. Остается показать, что любое N описанного выше вида порождает полугруппу T_t вероятностных операторов.

Для этой цели мы используем следующую лемму.

Лемма 4.2.1. *Рассмотрим последовательность полугрупп вероятностных операторов T_t^n, $n = 1, 2, \ldots$. Предположим, что их инфинитезимальные порождающие операторы существуют на C_2 и обозначаются $M^n f = A_n L_{g^{-1}}$ и что для любого $f \in C_2$ имеем $M^n f \to Mf$ в метрике C. Для соответствующих η-мер предположим, что $\eta_n(\omega) \to \eta(\omega)$. Тогда $T_t^n f$ сходится при $n \to \infty$ для каждой $f \in C$ к некоторому $T_t f$, где T_t есть полугруппа вероятностных операторов, инфинитезимальные порождающие операторы которых существуют на C_2 и совпадают с M.*

Доказательство здесь не приводится, так как оно громоздко, хотя и стандартно (см. замечания 4.2.3).

Пусть теперь M имеет вид

$$Mf(g) = \int\limits_{G_c - e} [f(gh) - f(g)]\, \eta(dh),$$

где η есть конечная мера на $G_c - e$. Но M — ограниченное преобразование C в C, и мы знаем из гл. 3, что $T_t = \exp(tM)$, $t \geqslant 0$, образуют полугруппу вероятностных операторов, непрерывных в равномерной операторной топологии, а M — ее инфинитезимальный оператор. Относительно вероятностной интерпретации этой полугруппы см. разд. 2.3.

Рассмотрим теперь M вида

$$Mf(g) = \int\limits_{G_c - e} \left[f(gh) - f(g) - \sum X_i f(g)\, x_i(h) \right] \eta(dh) + \sum a_i X_i f(g),$$

где η — по-прежнему конечная мера. Так как функции $x_i(h)$ ограничены, то можно писать

$$Mf(g) = Xf(g) + \int\limits_{G_c - e} [f(gh) - f(g)]\, \eta(dh),$$

где $X \in \Lambda(G)$. Но для каждого $f \in C_2$ можно аппроксимировать Xf интегральными операторами

$$\int\limits_{G_c - e} [f(gh) - f(g)]\, \lambda_n(dh) \to Xf(g)$$

в топологии C, где λ_n — конечные меры. Введем теперь

$$M^n f(g) = \int\limits_{G_c - e} [f(gh) - f(g)]\,[\eta(dh) + \lambda_n(dh)].$$

Приведенная выше лемма применима, и мы видим, что M есть инфинитезимальный порождающий оператор полугруппы вероятностных операторов.

Чтобы прийти к общему случаю, введем меры $\eta_n(dg) = [1 - \exp(-n\varphi(g))]\,\eta(dg)$. Тогда $\eta_n(dg)/\varphi(g)$ есть конечная мера на $G_c - e$ и

$$M^n f(g) = \sum a_i X_i f(g) + \int\limits_{G_c} [f(gh) - f(g)]\,\mu_n(dh) + \\ + \int\limits_{G_c - e} \left[f(gh) - f(g) - \sum X_i f(g)\, x_i(h) \right] \frac{\eta_n(dh)}{\varphi(h)},$$

где μ_n — конечная мера, такая, что

$$\int\limits_{G_c} [f(gh) - f(g)]\,\mu_n(dh) \to \sum a_{ij} X_i X_j f(g)$$

для любой $f \in C_2$ в топологии C. Но мы знаем, что M^n есть инфинитезимальный оператор некоторой полугруппы вероятностных операторов. Тогда из леммы вытекает, что M как предел M^n также есть инфинитезимальный порождающий оператор некоторой полугруппы вероятностных операторов.

Это приводит нас к следующей теореме относительно вида инфинитезимального порождающего оператора полугруппы вероятностных операторов.

Теорема 4.2.1. *Пусть T_t, $t \geqslant 0$, есть полугруппа вероятностных операторов на сепарабельной группе Ли. T_t имеет инфинитезимальный порождающий оператор M, определенный на C_2 соотношением*

$$Mf(g) = \sum a_i X_i f(g) + \sum a_{ij} X_i X_j f(g) + \\ + \int\limits_{G_c - e} \left[f(gh) - f(g) - \sum X_i f(g)\, x_i(h) \right] \frac{\eta(dh)}{\varphi(h)},$$

где a_i — действительные числа, a_{ij} образуют действительную, симметричную, неотрицательно определенную квадратную матрицу порядка d, а η — конечная мера *на G_c — e. Оператор M на C_2 определяет T_t. Обратно, если M определено указанным выше образом на C_2, то существует одна и только одна полугруппа вероятностных операторов T_t, имеющая M своим инфинитезимальным порождающим оператором в C_2.*

По аналогии со случаем действительной прямой введем следующее определение (см. замечания 4.2.$_4$).

Определение 4.2.1. *Вероятностная мера P на группе Ли называется нормальным (гауссовским) распределением, если существует полугруппа вероятностных операторов T_t с инфинитезимальным порождающим оператором M на C_2:*

$$Mf(g) = \sum a_i X_i f(g) + \sum a_{ij} X_i X_j f(g)$$

и 46

$$T_1 f(g) = \int_G f(gh) P(dh).$$

Вернемся на момент к компактным группам. Рассмотрим последовательность компактных групп

$$\Gamma_1, \ \Gamma_2, \ \Gamma_3, \ \ldots$$

и гомоморфизмы h_n, отображающие Γ_{n+1} на Γ_n. Введем множество G последовательностей $g = (\gamma_1, \gamma_2, \gamma_3, \ldots)$ таких, что $\gamma_n \in \Gamma_n$, $\gamma_n = h_n(\gamma_{n+1})$, и определим умножение двух последовательностей $g' = (\gamma_n')$, $g'' = (\gamma_n'')$ соотношением

$$g = g'g'' = (\gamma_1'\gamma_1'', \ \gamma_2'\gamma_2'', \ \ldots).$$

Затем в G вводится топология с помощью окрестностей N, где принадлежность N равносильна требованию, чтобы конечное число γ_n попадало в окрестности $N_n \subset \Gamma_n$. Тогда G есть компактная группа и называется *пределом групп* Γ_n.

Пусть теперь $P_t^{(n)}$ — непрерывные однородные процессы, определенные на Γ_n, $n = 1, 2, \ldots$ Нам необходимо согласованное определение, чтобы $P_t^{(n)}$ было распределением вероятностей в Γ_n, индуцированным распределением $P_t^{(n+1)}$ посредством гомоморфизма h_{n+1}. Тогда на G определен непрерывный однородный процесс.

Прежде всего, исходные распределения $P_t^{(n)}$ определяют распределение вероятностей P_t на G для каждого $t \geqslant 0$. Действительно, пусть $f(g)$ — непрерывная функция на G, зависящая только от $\gamma_1, \ldots, \gamma_n$. Тогда значение

$$Lf = \int\limits_G f(\gamma_1 \ldots \gamma_n) P_t^{(n)} (d\gamma_1 \ldots \gamma_n)$$

вполне определено. Используя компактность G и равномерную аппроксимацию такими функциями, можно непосредственно продолжить L до положительного функционала на $C(G)$. Это определяет регулярную вероятностную меру P_t на G. Однородность, $P_{t+s} = P_t * P_s$, устанавливается аналогичным образом. Наконец, имеет место непрерывность, так как если N есть произвольная окрестность e, содержащая $\gamma_1, \gamma_2, \ldots, \gamma_n$, то $P_t(N) = P_t^{(n)}(N) \to 1$ при $t \downarrow 0$, а это эквивалентно непрерывности (см. замечания $2.3._1$).

Заметим, между прочим, следующее. Если для всех n согласованным образом определены последовательности вероятностных мер $P_1^{(n)}, P_2^{(n)}, \ldots$ на Γ_n, то они определяют последовательность $P_1, P_2, \ldots$ на G. Если теперь для каждого n вероятностные меры $P_1^{(n)}, P_2^{(n)}, \ldots$ сходятся к $Q^{(n)}$, то это определяет $P \in \mathcal{P}(G)$ и $P_n \to P$. Это может быть доказано, как и выше, с помощью равномерной аппроксимации.

Вернемся теперь к однородным процессам. Заданная компактная группа G может быть представлена как предел групп Ли Γ_n (см. замечания $4.2._5$). Но мы знаем, какой вид могут иметь однородные процессы на группах Ли, т. е. мы знаем, как охарактеризовать их инфинитезимальные порождающие операторы (см. теорему 4.2.1). Для данного непрерывного однородного процесса P_t на G инфинитезимальный порождающий оператор M_n процесса $P_t^{(n)}$, индуцированного в Γ_n процессом P_t, имеет известную форму для любого n. Более интересно, что если оператор M сводится при любом n к оператору M_n указанного вида, если ограничиться Γ_n, то операторы M_n порождают множество процессов $P_t^{(n)}$, определенных согласованным образом, и, как мы видели выше, это приводит к непрерывному однородному процессу P_t на G.

Возможно, что при конкретном изучении распределений вероятностей на компактной группе удобно опираться на унитарные представления. В этом случае было бы полезно проследить связь с группами Ли.

4.3. Закон больших чисел на стохастических группах Ли

Рассмотрим треугольную систему распределений вероятностей на G_c

$$\begin{array}{lll} P_{11} & & \\ P_{21}, & P_{22} & \\ P_{31}, & P_{32}, & P_{33} \\ \dots & \dots & \dots \end{array}$$

В той или иной форме будет предполагаться, что эта система *инфинитезимальна*, т. е. что в каждой строке отдельные случайные величины пренебрежимы, как это делается в классической ситуации. Более точно, будем считать, что для любой окрестности N_0 элемента e должно иметь место равенство

$$\lim_{n\to\infty} \max_k P_{nk}(G_c - N_0) = 0.$$

Далее, мы будем предполагать, что в каждой строке $P_{ni} * P_{nj} = P_{nj} * P_{ni}$; в классических случаях это предположение коммутативности выполнено, конечно, автоматически. Заметим, что оно выполнено также тривиальным образом в важном случае *идентичных компонент*, т. е. в случае, когда $P_{ni} = P_{nj}$ для всех i и j.

Теорема 4.3.1. *Если треугольная система обладает указанными двумя свойствами и если*

$$1)\ \lim_{n\to\infty} \sum_{k=1}^{n} \left| \int_{N_0} x_i(g)\, P_{nk}(dg) \right| = 0;\ i = 1,\ 2,\ \dots,\ d,$$

$$2)\ \lim_{n\to\infty} \sum_{k=1}^{n} \int_{G_c} \varphi(g)\, P_{nk}(dg) = 0,$$

то $P_n = P_{n1} * P_{n2} * \dots * P_{nn}$ *сходится к вырожденному распределению* δ_e.

Доказательство. Достаточно доказать сильную сходимость $T_n = T_{n1}T_{n2}\ldots T_{nn} \to I$. Но так как $\|T_n - I\| \leqslant 2$, достаточно показать, что $\|(T_{nk} - I) f\| \to 0$ при $f \in C_2$, так как C_2 плотно в C.

Рассмотрим

$$[T_{nk} - I] f(g) = \int\limits_{G_c - N_0} [f(gh) - f(g)] P_{nk}(dh) + \int\limits_{N_0} [f(gh) - f(g)] P_{nk}(dh).$$

С первым слагаемым дело обстоит просто, так как

$$\left| \int\limits_{G_c - N_0} [f(gh) - f(g)] P_{nk}(dh) \right| \leqslant 2 \|f\| P_{nk}(G_c - N_0),$$

и из условия 2) вытекает, что

$$\lim_{n\to\infty} \sum_{k=1}^{n} P_{nk}(G_c - N_0) = 0.$$

Чтобы мажорировать второе слагаемое, разлагаем подинтегральную функцию в ряд Тейлора в окрестности N_0:

$$f(gh) - f(g) = \sum_i X_i f(g)\, x_i(h) + \frac{1}{2} \sum_{i,j} X_i X_j f(g\theta)\, x_i(h)\, x_j(h).$$

Но тогда

$$\left| \int\limits_{N_0} [f(gh) - f(g)] P_{nk}(dh) \right| \leqslant \sum_i |X_i f(g)| \cdot \left| \int\limits_{N_0} x_i(h) P_{nk}(dh) \right| + \frac{1}{2} \sum_{i,j} \int\limits_{N_0} |X_i X_j f(g\theta)| \cdot |x_i(h)| \cdot |x_j(h)| P_{nk}(dh) \leqslant$$
$$\leqslant \|f\|_2 \sum \left| \int\limits_{N_0} x_i(h) P_{nk}(dh) \right| + \text{const} \cdot \|f\|_2 \int\limits_{N_0} \varphi(h)\, P_{nk}(dh).$$

Отсюда, учитывая условия 1) и 2), получаем

$$\lim_{n\to\infty} \sum \|[T_{nk} - I] f\| = 0.$$

Но

$$[T_n - I] f = [T_{n1} T_{n2} \ldots T_{nn} - I] f = [T_{n1} - I] f + \\ + [T_{n1} T_{n2} - T_{n1}] f + \ldots + [T_{n1} T_{n2} \ldots T_{nn} - T_{n1} T_{n2} \ldots T_{nn-1}] f,$$

так что

$$\| [T_n - I] f \| \leqslant \sum_1^n \| T_{n1} T_{n2} \ldots T_{nk-1} (T_{nk} - I) f \| \leqslant \\ \leqslant \sum_1^n \| [T_{nk} - I] f \| \to 0,$$

что завершает доказательство.

На действительной прямой закон больших чисел выполнен при любом среднем значении, что вытекает из коммутативности группы действительных чисел по сложению. Здесь можно было бы только выдвинуть предположение, что при замене условия 1) условием

$$\lim_{n \to \infty} \sum_{k=1}^n \int_{N_0} x_i (g) P_{nk} (dg) = a_i; \ \ i = 1, 2, \ldots, d,$$

распределенная масса сходится к элементу $\exp A$, где $A = \Sigma a_i X_i \in \Lambda (G)$. Доказать это непосредственным обобщением предыдущего доказательства не представляется возможным, так как отсутствие коммутативности вызывает некоторые трудности. Поэтому мы должны отложить обсуждение этой проблемы.

4.4. Центральная предельная теорема

Если в произведении $T_n = T_{n1} T_{n2} \ldots T_{nn}$ или $P_{n1} * P_{n2} * \ldots * P_{nn}$ множители близки друг к другу, то можно ожидать, что $T_{nk} \cong \exp (T_{nk} - I)$. В этом случае можно также рассчитывать на то, что эти множители приближенно перестановочны, так что

$$V_n = \exp \sum_{k=1}^n (T_{nk} - I) \cong T_n.$$

Соотношение этого типа существенно при изучении предельных теорем на группах Ли. Докажем следующую теорему.

Теорема 4.4.1. *Пусть вероятностные меры P_{nk} образуют инфинитезимальную и коммутативную треугольную систему, такую, что*

1) $\sum_{k=1}^{n} \left| \int_{N_0} x_i(g) P_{nk}(dg) \right|$ *равномерно ограничены,*

2) $\sum_{k=1}^{n} \int_{G_c} \varphi(g) P_{nk}(dg)$ *равномерно ограничены,*

3) *для любого положительного* ε *существует такое компактное множество* $F = F_\varepsilon$, *что для всех* n *имеет место неравенство*

$$\sum_{k=1}^{n} P_{nk}(G-F) < \varepsilon.$$

Тогда $(T_n - V_n) f \to 0$ *для всех* $f \in C$ *по норме в* C.

Доказательство. При $S_{nk} = \exp(T_{nk} - I)$ имеем, в силу того что при фиксированном n все T_{nk} и S_{nk} коммутируют,

$$(T_n - V_n) f = \sum_{k=1}^{n} T_{n1} \ldots T_{nk-1} S_{nk+1} \ldots S_{nn} (T_{nk} - S_{nk}) f,$$

что показывается, как в конце предыдущего раздела. Следовательно,

$$\| (T_n - V_n) f \| \leqslant \sum_{k=1}^{n} \| (T_{nk} - S_{nk}) f \|,$$

так как S_{nk} есть (пуассоновский) вероятностный оператор, $\|S_{nk}\| = 1$. Но

$$T_{nk} - S_{nk} = T_{nk} - \left\{ I + (T_{nk} - I) + \frac{(T_{nk} - I)^2}{2!} + \ldots \right\} =$$
$$= -A_{nk} (T_{nk} - I)^2,$$

где

$$A_{nk} = \frac{1}{2!} I + \frac{(T_{nk} - I)}{3!} + \frac{(T_{nk} - I)^2}{4!} + \ldots,$$

так что A_{nk} являются операторами с равномерно ограниченной нормой. Следовательно,

$$\| (T_n - V_n) f \| \leqslant \mathrm{const} \cdot \sum_{k=1}^{n} \| (T_{nk} - I)^2 f \|.$$

Достаточно доказать, что правая часть этого неравенства стремится к нулю при $n \to \infty$ для всех $f \in C_2$, так как C_2 плотно в C.

Оставшаяся часть доказательства принципиально не представляет затруднений, хотя подробное его проведение требует внушительной вычислительной работы. Чтобы дать читателю представление о существе доказательства, предположим, что N есть окрестность элемента e, настолько малая, что имеют место приведенные ниже разложения в ряды Тейлора с установленными остаточными членами. Запишем

$$(T_{nk} - I)^2 f(g) =$$
$$= \int_{G_c} \left\{ \int_{G_c} [f(gxy) - f(gx) - f(gy) + f(g)] P_{nk}(dy) \right\} P_{nk}(dx) =$$
$$= \int_{G_c - N} \int_{G_c - N} + \int_{G_c - N} \int_{N} + \int_{N} \int_{G_c - N} + \int_{N} \int_{N} = I_{1k} + I_{2k} + I_{3k} + I_{4k}.$$

Для оценки I_{1k} заметим, что

$$| I_{1k} | \leqslant 4 \| f \| P_{nk}^2 (G_c - N),$$

так что

$$\sum_{k=1}^{n} | I_{1k} | \leqslant 4 \| f \| \sum_{k=1}^{n} P_{nk}(G_c - N) \cdot \max P_{nk}(G_c - N) \to 0$$

в силу условия 2) и того, что система $\{P_{nk}\}$ инфинитезимальна.

Для I_{2k} напишем

$$\int_{N} [f(gxy) - f(gx) - f(gy) - f(g)] P_{nk}(dy) =$$
$$= \int_{N} [f(gxy) - f(gx)] P_{nk}(dy) - \int_{N} [f(gy) - f(g)] P_{nk}(dy)$$

и получим, применяя теорему Тейлора,

$$\int_N [f(gh)-f(g)]\,P_{nk}(dh)=\sum_i X_i f(g)\int_N x_i(h)\,P_{nk}(dh)+\\+\frac{1}{2}\int_N \sum_{i,j} X_i X_j f(g\theta)\,x_i(h)\,x_j(h)\,P_{nk}(dh),$$

так что

$$\left|\int_N [f(gh)-f(g)]\,P_{nk}(dh)\right|\leqslant \|f\|_1\cdot\sum_i\left|\int_N x_i(h)\,P_{nk}(dh)\right|+\\+\mathrm{const}\cdot\|f\|_2\int_N \varphi(h)\,P_{nk}(dh).$$

Следовательно, суммируя такие члены, получим

$$\sum_{k=1}^{n}|I_{2k}|\leqslant 2\|f\|_2\sum_{k=1}^{n}\Big[\sum_i\left|\int_N x_i(h)\,P_{nk}(dh)\right|+\\+\mathrm{const}\cdot\int_N \varphi(h)\,P_{nk}(dh)\Big]\,P_{nk}(G_c-N),$$

что стремится к нулю при $n\to\infty$.

В четвертом интеграле обе переменные интеграции x и y изменяются в малой окрестности N, и разложение в ряд Тейлора и кропотливые выкладки, аналогичные проведенным выше, показывают, что вклад этого интеграла мал при больших n. Третий интеграл

$$I_{3k}=\int_{G_c-N}\left\{\int_N [f(gxy)-f(gy)-f(gx)+f(g)]\,P_{nk}(dx)\right\}\,P_{nk}(dy)$$

более интересен.

При изучении функции

$$f(gxy)-f(gy)=f[gy(y^{-1}xy)]-f(gy),$$

где x меняется в некоторой малой окрестности элемента e, мы не можем быть уверены, что $y^{-1}xy$ изменяется в некоторой малой и фиксированной окрестности элемента e для всех $y\in G_c-N$. Из-за этой трудности введено условие 3).

Запишем

$$I_{3k}=\left(\int\limits_{G-F}+\int\limits_{F-N}\right)\left\{\int\limits_{N}[f(gxy)-f(gx)-f(gy)+f(g)]\times\right.$$

$$\left.\times P_{nk}(dx)\right\}P_{nk}(dy)+\int\limits_{N}[f(g)-f(gx)]\,P_{nk}(dx)\,P_{nk}(\omega),$$

где последний член есть вклад элемента $y=\omega$. Первый интеграл не представляет затруднений, так как

$$\sum_{k=1}^{n}\left|\int\limits_{G-F}\int\limits_{N}\right|\leqslant 4\,\|f\|\sum_{k=1}^{n}P_{nk}(G-F)\leqslant 4\,\|f\|\,\varepsilon$$

в соответствии с тем, как было введено множество F в формулировке теоремы. Интегралы

$$\int\limits_{N}[f(g)-f(gx)]\,P_{nk}(dx)\,P_{nk}(\omega)$$

и

$$\int\limits_{F-N}\left\{\int\limits_{N}[f(gx)-f(g)]\,P_{nk}(dx)\right\}P_{nk}(dy)$$

снова подвергаются разложениям в ряды. Остается член

$$\int\limits_{F-N}\left\{\int\limits_{N}[f(gxy)-f(gy)]\,P_{nk}(dx)\right\}P_{nk}(dy).$$

Из разложения в ряд имеем

$$\left|\int\limits_{N}[f(gxy)-f(gy)]\,P_{nk}(dx)\right|\leqslant\|f\|_1\sum_{i}\left|\int\limits_{N}x_i(y^{-1}xy)\,P_{nk}(dx)\right|+$$

$$+\operatorname{const}\cdot\|f\|_2\int\limits_{N}\varphi(y^{-1}xy)\,P_{nk}(dx).$$

Решающим обстоятельством является то, что y меняется только на компактном множестве. Максимумы членов в правой части достигаются тогда для некоторых значений $y_1,\ \ldots,\ y_d$ и y_0, так что

$$\sum_{k=1}^{n}\left|\int\limits_{F-N}\int\limits_{N}\right|\leqslant\|f\|_2\sum_{k=1}^{n}\left\{\sum_{i}\left|\int\limits_{N}x_i(y_i^{-1}xy_i)\,P_{nk}(dx)\right|+\right.$$

$$\left.+\operatorname{const}\cdot\int\limits_{N}\varphi(y_0^{-1}xy_0)\,P_{nk}(dx)\right\}P_{nk}(F-N)\to 0$$

при $n \to \infty$, так как система является инфинитезимальной. Это завершает доказательство теоремы.

Теперь мы приходим к одному из вариантов центральной предельной теоремы.

Теорема 4.4.2. *Пусть* $\{P_{nk}\}$ — *коммутативная инфинитезимальная система распределений вероятностей на группе Ли, такая, что*

1) $\lim\limits_{n\to\infty} \sum\limits_{k=1}^{n} \int\limits_{N_0} x_i(g) P_{nk}(dg) = a_i$,

$$\sum_{k=1}^{n} \left| \int\limits_{N_0} x_i(g) P_{nk}(dg) \right| \text{ *равномерно ограничены*,}$$

2) $\lim\limits_{n\to\infty} \sum\limits_{k=1}^{n} \int\limits_{N_0} x_i(g)\, x_j(g)\, P_{nk}(dg) = a_{ij}$,

3) $\lim\limits_{n\to\infty} \sum\limits_{k=1}^{n} \int\limits_{G_C - N_0} \varphi(g)\, P_{nk}(dg) = 0$

для каждой окрестности N_0 *элемента* e.

Тогда $P_n = P_{n1} * P_{n2} * \ldots * P_{nn}$ *сходится к нормальному распределению, порожденному оператором*

$$M = \sum_i a_i X_i + \sum_{i,\,j} a_{ij} X_i X_j.$$

Доказательство. Используя теорему 4.4.1, достаточно только рассмотреть $P_n = \exp \sum\limits_{1}^{n} (T_{nk} - I)$ и его предел при $n \to \infty$. Но P_n может быть вложено в полугруппу $\exp t \sum\limits_{1}^{n} (T_{nk} - I)$ сложного пуассоновского типа, и его инфинитезимальный порождающий оператор может быть записан в виде ($f \in C_2$)

$$\sum_{k=1}^{n} (T_{nk}-I) f(g) = \sum_{k=1}^{n} \int\limits_{G_C} [f(gh)-f(g)]\, P_{nk}(dh) =$$

$$= \sum_{i} a_i^{(n)} X_i f(g) + \int\limits_{G_C - e} \Big[f(gh) - f(g) - \sum_{i} X_i f(g)\, x_i(h) \Big] \frac{\eta_n(dh)}{\varphi(h)},$$

где

$$\eta_n(dh) = \sum_{k=1}^{n} \varphi(h)\, P_{nk}(dh),$$

$$a_i^{(n)} = \int\limits_{N_0} x_i(h) \frac{\eta_n(dh)}{\varphi(h)} + \int\limits_{G_C - N_0} x_i(h) \frac{\eta_n(dh)}{\varphi(h)}.$$

Используем теперь условие 1), чтобы показать, что первый член в $a_i^{(n)}$ стремится к a_i, и условие 3), чтобы показать, что второй член стремится к нулю. Далее, η_n стремится к нулю вне любой окрестности элемента e. Отсюда следует, что имеет место сильная сходимость

$$\sum_{k=1}^{n} (T_{nk}-I) f \to \sum_{i} a_i X_i f + \sum_{i,\,j} a_{ij} X_i X_j f,$$

а, значит, (см. замечания 4.2.$_3$ и разд. 4.2) P_n и T_n сходятся к $\exp\{\sum_i a_i X_i + \sum_{i,j} a_{ij} X_i X_j\}$, что и утверждалось.

Вернемся теперь на некоторое время к закону больших чисел. Перефразируем теорему 4.3.1 следующим образом.

Т е о р е м а 4.3.1.*a*. *Пусть* $\{P_{nk}\}$ *есть инфинитезимальная и коммутативная система распределений вероятностей, такая, что*

1) $\lim\limits_{n\to\infty} \sum\limits_{k=1}^{n} \int\limits_{N_0} x_i(g)\, P_{nk}(dg) = a_i,$

$$\sum_{k=1}^{n} \Big| \int\limits_{N_0} x_i(g)\, P_{nk}(dg) \Big| \quad \textit{равномерно ограничены,}$$

2) $\lim\limits_{n\to\infty} \sum\limits_{k=1}^{n} \int\limits_{G_C} \varphi(g)\, P_{nk}(dg) = 0.$

Тогда $P_n = P_{n1} * P_{n2} * \ldots * P_{nn}$ *сходится к вырожденному распределению, вся масса которого сосредоточена в точке* $\exp \sum_i a_i X_i \in G$.

Доказательство. Это сводится к последней теореме, но с $a_{ij} = 0$, так что P_n и T_n стремятся к $\exp \sum_i a_i X_i$ — вероятностному оператору, соответствующему распределению вероятностей с массой, сосредоточенной в точке $\exp \sum_i a_i X_i \in G$.

4.5. Примеры

4.5.1. Простейшим броуновским движением является движение по прямой, $x(t) = N(0, \sigma^2 t)$, в предположении отсутствия сноса. Его функция плотности

$$f(x, t) = \frac{1}{\sigma \sqrt{2\pi t}} \exp\left[-\frac{x^2}{2\sigma^2 t}\right]$$

удовлетворяет обычному уравнению теплопроводности

$$\frac{\partial f}{\partial t} = \frac{\sigma^2}{2} \frac{\partial^2 f}{\partial x^2}$$

с начальным условием: при $t = 0$ вся масса в точке $x = 0$. На одномерном торе имеем то же самое дифференциальное уравнение, но с идентификацией x_1 и x_2 при $x_1 \equiv x_2 \pmod{2\pi}$. Другими словами, мы изучаем это уравнение на интервале $(-\pi, \pi)$ с граничным условием $f(-\pi, t) = f(\pi, t)$, а эта задача имеет хорошо известное решение

$$f(x, t) = \frac{1}{\sigma \sqrt{2\pi t}} \sum_{\nu=-\infty}^{\infty} \exp\left[-\frac{(x - 2\nu\pi)^2}{2\sigma^2 t}\right].$$

Этот ряд является просто суперпозицией предыдущего решения в точках, конгруэнтных с x по модулю 2π. Заметим, что R^1 есть накрывающая группа T^1.

Этот простой факт поддается значительному обобщению. Пусть G — связная, локально односвязная группа, предполагаемая, как обычно, локально компактной и сепарабельной. В данном контексте можно представлять себе сепарабельную связную группу Ли. Пусть G^* есть *накры-*

вающая группа для G с гомоморфизмом φ группы G^* на G. Обозначим через $N = \{g_1^*, g_2^*, \ldots\}$ счетный нормальный делитель группы G^*, такой, что G изоморфна G^*/N, так что образом $g \in G$ является $g_1^* g^*$, $g_2^* g^*$, ... Разумно выбрать G^* как универсальную накрывающую группу для G; если бы нас интересовала любая другая накрывающая группа Γ для G, то Γ накрывалась бы G^*.

Начнем с броуновского движения $g(t)$, $t \geqslant 0$, на G (см. 3.1). Мы определяем стохастический процесс на G^* выборочной функцией, следуя, исходя из элемента e, за элементом из G^*, соответствующим $g(t)$ для $0 \leqslant \tau \leqslant t$, и обозначая конечную точку этой кривой через $g^*(t)$. Тогда $g^*(t)$ есть также броуновское движение с вероятностной мерой P_t^* на G^*. Для полного доказательства этого утверждения потребуются некоторые рассуждения, относящиеся к теории меры, но мы предоставим их терпеливому читателю. Более интересно следующее. Пусть S есть борелевское множество из G^*. Тогда $P_t^*(S) = 0$ тогда и только тогда, когда $P_t(\varphi S) = 0$, так как N счетно. Пусть G имеет меру Хаара dg и G^* имеет меру Хаара dg^*. Пусть P_t^* абсолютно непрерывна относительно dg^*:

$$P_t^*(E) = \int\limits_E p^*(g^*, t)\, dg^*, \quad E \subset G^*.$$

Для заданного $g_0 \in G$ выберем окрестность N^* элемента e^* столь малую, что $g_n^* g_0^* N^*$ не имеют общих элементов для разных n (g_0^* есть некоторый прообраз g_0), и такую, что N^* и $\varphi(N^*) = N$ изоморфны. Если E — борелевское множество в $g_0 N$, то $\varphi^{-1}E$ есть сумма непересекающихся множеств $g_n^*[\varphi^{-1}E \cap g_0^* N^*]$, так что

$$P_t(E) = \sum_1^\infty \int\limits_{g_n^*[\varphi^{-1}E \cap g_0^* N^*]} p^*(t, g^*)\, dg^* =$$

$$= \sum_1^\infty \int\limits_{\varphi^{-1}E \cap g_0^* N^*} p^*(t, g_n^* g^*)\, dg^* =$$

$$= \int\limits_{\varphi^{-1}E \cap g_0^* N^*} \sum_1^\infty p^*(t, g_n^* g^*)\, dg^*.$$

Так как сумма под знаком интеграла не зависит от того, какой образ g^* выбирается для g, и так как мера Хаара dg^* на $g_n^* N^*$ может быть идентифицирована с мерой Хаара dg на $g_0 N$, то

$$P_t(E) = \int\limits_E \sum_1^\infty p^*(t, g_n^* g^*)\, dg = \int\limits_E p(t, g)\, dg,$$

где

$$p(t, g) = \sum_1^\infty p^*(t, g_n^* g^*)$$

почти наверное. Это и есть желаемое обобщение. Мы использовали здесь инвариантные слева меры Хаара.

Пусть P — нормальное распределение на T^1. Тогда P может быть записано как композиция $P_1 * P_2$, где P_1 и P_2 не являются нормальными (см. замечания 4.5.1). Можно сказать, что теорема Крамера не имеет места на T^1. Отсюда вытекает ряд интересных следствий. Представим себе центральную предельную теорему на прямой. Мы знаем, что для того, чтобы она имела место, нужно, чтобы каждое слагаемое в сумме было мало или имело приближенно нормальное распределение. На T^1 при больших и существенно не нормальных слагаемых можно все же прийти к нормальному распределению.

4.5.2. Рассмотрим теперь двумерную группу Ли, состоящую из квадратных матриц второго порядка

$$g = \begin{Bmatrix} 1 & g_1 \\ 0 & g_2 \end{Bmatrix}, \quad g_2 > 0,$$

— *аффинную группу на действительной прямой*. Этот простой, но показательный пример будет изучаться с другой точки зрения в 5.5.1 (см. также 7.5.1). Для того чтобы найти дифференциальный оператор броуновского движения на этой группе, удобно получить инфинитезимальные моменты первого и второго порядка и выписать соответствующее уравнение Фоккера — Планка. Тогда получим

$$g(t+h) = g(t) \begin{Bmatrix} 1 & \varepsilon_1(h) \\ 0 & 1 + \varepsilon_2(h) \end{Bmatrix},$$

где функции ε можно трактовать как нормальные случайные величины при малых $h > 0$. Это эквивалентно

$$\begin{cases} g_1(t+h) = \varepsilon_1(h) + g_1(t)\,[1+\varepsilon_2(h)] \\ g_2(t+h) = g_2(t)\,[1+\varepsilon_2(h)] \end{cases}$$

или

$$\begin{cases} g_1(t+h) - g_1(t) = \varepsilon_1(h) + g_1(t)\,\varepsilon_2(h) \\ g_2(t+h) - g_2(t) = g_2(t)\,\varepsilon_2(h). \end{cases}$$

Инфинитезимальные моменты должны тогда иметь вид

$$\lim_{h \downarrow 0} \mathbf{E}\,\frac{\Delta g_1}{h} = a_1 + a_2 g_1,$$

$$\lim_{h \downarrow 0} \mathbf{E}\,\frac{\Delta g_2}{h} = a_2 g_2,$$

$$\lim_{h \downarrow 0} \mathbf{E}\,\frac{(\Delta g_1)^2}{h} = b_1 + b_2 g_1^2 + 2b_3 g_1,$$

$$\lim_{h \downarrow 0} \mathbf{E}\,\frac{(\Delta g_2)^2}{h} = b_2 g_2^2,$$

$$\lim_{h \downarrow 0} \mathbf{E}\,\frac{\Delta g_1 \Delta g_2}{h} = g_1 g_2 b_2 + g_2 b_3.$$

Для удобства выделим случай $b_1 = b_2 = 1$, $b_3 = 0$, приводящий к уравнению

$$\frac{\partial p}{\partial t} = \frac{1}{2}\,\frac{\partial^2}{\partial g_1^2}\,[(1+g_1^2)\,p] + \frac{1}{2}\,\frac{\partial^2}{\partial g_2^2}\,[g_2^2 p] + \\ + \frac{\partial^2}{\partial g_1\,\partial g_2}\,[g_1 g_2 p] - \frac{\partial}{\partial g_1}\,[(a_1 + a_2 g_1)\,p] - \frac{\partial}{\partial g_2}\,[a_2 g_2 p].$$

Для изучения частных распределений g_1 и g_2 нужно произвести интегрирование соответственно по g_2 и g_1. Частная плотность p_2 величины g_2 удовлетворяет тогда уравнению

$$\frac{\partial p_2}{\partial t} = \frac{1}{2}\,\frac{\partial^2}{\partial g_2^2}\,(g_2^2 p_2) - \frac{\partial}{\partial g_2}\,(a_2 g_2 p_2).$$

Это уравнение имеет решение

$$p_2(g_2) = \frac{1}{\sqrt{2\pi t}\,g_2}\exp\left[-\frac{\log g_2 - \left(a_2 - \frac{1}{2}\,t\right)^2}{2t}\right],$$

что не удивительно, так как это плотность логонормального распределения, что естественно ожидать. С точки зрения g_2-координат групповая операция есть просто умножение, и так как мы должны иметь дело с большим числом независимых сомножителей, каждый из которых близок к единице, то естественно ожидать логонормального распределения. Частная плотность p_1 величины g_1 удовлетворяет уравнению

$$\frac{\partial p_1}{\partial t}=\frac{1}{2}\frac{\partial^2}{\partial g_1^2}[(1+g_1^2)p_2]-\frac{\partial}{\partial g_1}[(a_1+a_2g_1)p_2].$$

Интересно, что может существовать стационарное распределение для g_1 (но, конечно, не для g_2 за тривиальным исключением). Такая стационарная плотность должна удовлетворять уравнению

$$\frac{1}{2}\frac{d^2}{dg_1^2}[(1+g_1^2)p_1]-\frac{d}{dg_1}[(a_1+a_2g_1)p_1]=0,$$

так что

$$\frac{1}{2}\frac{d}{dg_1}[(1+g_1^2)p_1]-(a_1+a_2g_1)p_1=\frac{1}{2}(1+g_1^2)p_1'-$$
$$-[a_1+(a_2-1)g_1]p_1=C_1,$$

с общим решением

$$p_1(g_1)=(1+g_1^2)^{a_2-1}\exp(2a_1\operatorname{arctg}g_1)\times$$
$$\times\left[2C_1\int_0^{g_1}(1+h^2)^{-a_2}\exp(-2a_1\operatorname{arctg}h)\,du+C_2\right].$$

При $a_2\geqslant{}^1/_2$ эта функция неинтегрируема на $(-\infty,\infty)$ и, следовательно, не является плотностью. При $a_2<{}^1/_2$, интеграл внутри квадратных скобок ведет себя как $g_1^{-2a_2+1}$, так что, для того чтобы получить плотность, нужно положить $C_1=0$. Тогда мы получим стационарную плотность

$$p_1(g_1)=C_2(1+g_1^2)^{a_2-1}\exp(2a_1\operatorname{arctg}g_1).$$

Следствия из этого будут изучены позднее.

4.5.3. Рассмотрим группу *G дробно-линейных преобразований*

$$z_1\circ z=\frac{z_1+z}{1+\bar{z}_1z},$$

оставляющих инвариантным единичный открытый круг $C=\{z \mid |z|<1\}$. Читатель, знакомый с неевклидовой геометрией, узнает здесь параллельные движения на плоскости Лобачевского в круговой модели. Предположим теперь, что в групповом произведении $z_1 \circ z_2 \circ \ldots \circ z_n$ множители независимы и все имеют одинаковое распределение вероятностей $P_{nk}=P_n$ и что P_n инвариантны по отношению к вращениям (изотропны) в C; тогда можно положить $P_{nk} \to \delta_e$ и применить теорему 4.4.2.

Это приводит к броуновскому движению на C, соответствующему оператору Лапласа. В этом специальном случае плотность может быть получена в явной форме (см. замечания 4.5.3).

Глава 5

ЛОКАЛЬНО КОМПАКТНЫЕ СТОХАСТИЧЕСКИЕ ГРУППЫ

5.1. Унитарные представления

Для двух типов групп, исследованных в гл. 3, мы существенно опирались на компактность или предположение коммутативности, а в гл. 4 мы использовали дифференцируемость. Теперь, когда мы предполагаем группу локально компактной, мы можем ожидать больших трудностей, и в то же время сталкиваемся с новыми задачами, бросающими вызов вероятностной интуиции. Одной из главных задач общей теории стохастических групп является создание связной картины вероятностного поведения локально компактных групп. В настоящее время наши знания в этом направлении имеют серьезные пробелы, и очевидность этого будет болезненно чувствоваться читателем при чтении следующей главы.

Напомним сначала читателю некоторые основные факты относительно представлений, положительно определенных функций и т. п. для локально компактных групп. Как и раньше, мы будем рассматривать только сепарабельные группы; это ограничение может быть не очень существенным и делается главным образом из соображений удобства. Пусть G такая группа. Рассмотрим *унитарное представление* $r = (\mathscr{H}, U(g))$, где $\mathscr{H}$ — сепарабельное гильбертово пространство, а $U(g)$ — унитарное преобразование пространства $\mathscr{H}$ при любом $g \in G$. Мы требуем, чтобы для любого $z \in \mathscr{H}$ функция $U(g)z$, принимающая векторные значения, была сильно непрерывной (заметим, что здесь это эквивалентно слабой непрерывности) и чтобы $U(g)$ удовлетворяла уравнению $U(g_1)\ U(g_2) = U(g_1 g_2)$ для всех $g_1, g_2 \in G$. Если не существует нетривиального замкнутого подпространства пространства $\mathscr{H}$, инвариантного слева для всех $U(g)$, $g \in G$, то r называется *неприводимым*. Так же, как и для компактных групп (см. разд. 3.1), мы, по существу, изучаем классы эквивалентности. Два различных класса состоят из неэквивалентных представлений. Заметим, что для компактных групп всегда можно выбрать в качестве пространства $\mathcal{H}$, соответствующего неприводимому унитарному представлению, конечномерное унитарное пространство, а для коммутативных групп даже одномерное пространство.

В качестве примера пусть $\mathcal{H}$ будет гильбертовым пространством всех функций $f(h)$ с интегрируемым квадратом относительно инвариантной слева меры Хаара. Определим $U(g)$ соотношением $U(g)f(h) = L_g f(h) = f(g^{-1}h)$. Тогда $U(g)$ — унитарное (но не обязательно неприводимое) представление и называется левым *регулярным представлением*.

Пусть $\{\mathcal{H}, U(g)\}$ — унитарное представление группы G. Введем комплекснозначную функцию $\varphi(g) = (U(g)z, z)$ для некоторого $z \in \mathcal{H}$. Тогда для любого n и любых комплексных чисел $c_1, c_2, \ldots, c_n$ имеем

$$\sum_{\nu,\,\mu=1}^{n} c_\nu \bar{c}_\mu \varphi(g_\mu^{-1} g_\nu) = \left\| \sum_{\nu=1}^{n} c_\nu U(g_\nu) z \right\|^2 \geqslant 0.$$

По аналогии с действительной прямой функцию, удовлетворяющую такому неравенству, будем называть *положительно определенной* на G. Мы будем рассматривать только непрерывные положительно определенные функции (вроде той, которая была определена выше). Обратно, всякая заданная непрерывная и положительно определенная функция может быть представлена как $(U(g)z, z)$ при соответствующем выборе $(\mathcal{H}, U(g))$ и z. Это соответствие между положительно определенными функциями и унитарными представлениями многократно будет использоваться в дальнейшем. Если $\varphi(g) \not\equiv 0$, то $\varphi(e) > 0$, так что мы можем нормировать положительно определенные функции, требуя, чтобы $\varphi(e) = 1$, что и будет иногда делаться.

Ясно, что множество Φ положительно определенных функций выпукло: $c\varphi_1 + (1-c)\varphi_2 \in \Phi$ для всех $c \in (0, 1)$, если $\varphi_1, \varphi_2 \in \Phi$. Среди функций из Φ мы следующим образом выберем экстремальную. Сделаем множество Φ частично упорядоченным, положив $\varphi_1 \prec \varphi_2$, если $\varphi_2 - \varphi_1 \in \Phi$, и будем говорить в этом случае, что φ_1 подчинено φ_2. Положительно определенная функция φ, для которой единственными подчиненными функциями будут кратные $c\varphi$, $c \in (0, 1)$, называется *элементарной положительно определенной функцией*. Установлено, что такие функции в точности соответствуют неприводимым унитарным представлениям.

Некоторые, хотя и не все, положительно определенные функции могут быть получены следующим образом с помощью композиции. Пусть $a(g)$ есть функция из $L(G)$, равная нулю вне некоторого компактного множества C. Образуем функцию

$$p(g) = \int_G a(h)\bar{a}(hg)\,\nu(dh).$$

Тогда $p(g)$ непрерывна и положительно определена, так как

$$\sum_{i,\,j=1}^{n} c_i\bar{c}_j p(g_i g_j^{-1}) = \sum_{i,\,j=1}^{n} c_i\bar{c}_j \int_G a(h)\bar{a}(hg_i g_j^{-1})\,\nu(dh) =$$

$$= \sum_{i,\,j=1}^{n} c_i\bar{c}_j \int_G a(kg_i^{-1})\bar{a}(kg_j^{-1})\,\nu(dk) =$$

$$= \int_G \left|\sum_{i=1}^{n} c_i a(kg_i^{-1})\right|^2 \nu(dk) \geqslant 0.$$

Заметим также, что $p(g)$ равно нулю, если $g \bar{\in} C^{-1}C$. Последнее множество компактно (см. замечания 5.1.$_2$).

Основной результат заключается в том, что существует множество $R = \{r\}$ неприводимых унитарных представлений, *полное* в следующем смысле: если g_0—произвольный элемент из G, отличный от единичного, то существует такое $r \in R$, что $U_r(g_0) \neq I$. Другими словами, неприводимых унитарных представлений достаточно много, чтобы разделять точки из G.

Из свойства полноты вытекают важные следствия. Можно рассматривать элементарные положительно определенные функции как «блоки» для построения конечных линейны комбинаций $c_1\varphi_1 + c_2\varphi_2 + \ldots + c_n\varphi_n$, так называемых *тригонометрических полиномов*. Это обстоятельство будет для нас полезным, и мы сформулируем его тремя различными способами.

а) Если μ — ограниченная комплексная мера и

$$\int\limits_{g \in G} \varphi(g)\, \mu(dg) = 0$$

для всех элементарных положительно определенных функций φ, то $\mu = 0$.

б) Любая непрерывная функция на G может быть аппроксимирована тригонометрическими полиномами равномерно на каждом компактном множестве.

в) Любая непрерывная положительно определенная функция φ_0 может быть равномерно аппроксимирована на G функцией ψ вида

$$\psi(g) = \int\limits_{\varphi \in E} \varphi(g)\, \nu(dg),$$

где E — компактное множество элементарных положительно определенных функций φ, а ν — борелевская мера на E (см. замечания $5.1._3$).

Уже несколько десятилетий известно, что в компактном случае любое унитарное представление может быть разложено в прямую сумму неприводимых унитарных представлений. В нашем случае эта дискретная сумма должна быть заменена интегралом. Пусть D — множество элементов δ с мерой ϱ. Каждому $\delta \in D$ соответствует гильбертово пространство $\mathscr{H}_\delta$ со скалярным произведением $(u, v)_\delta$. Образуем теперь новое гильбертово пространство

$$\mathscr{H} = \int\limits_{\delta \in D} \mathscr{H}_\delta \sqrt{\varrho(d\delta)},$$

состоящее из функций $x = x_\delta$, $x_\delta \in \mathscr{H}_\delta$ для любого $\delta \in D$, с конечной нормой, соответствующей скалярному произведению

$$(x, y) = \int\limits_{\delta \in D} (x_\delta, y_\delta)_\delta \varrho(d\delta).$$

Назовем $\mathscr{H}$ *прямым интегралом* пространств $\mathscr{H}_\delta$. Можно доказать, что любое унитарное представление эквивалентно представлению в пространстве $\mathscr{H}$, являющемся прямым произведением пространств $\mathscr{H}_\delta$; кроме того, $U(g)x = \{U_\delta(g)\, x_\delta,\ \delta \in D\}$, где $U_\delta(g)$ — неприводимое унитарное преобразование для почти всех δ (по мере ϱ). Функции $(U_\delta(g)x, y)$, зависящие от g и δ, измеримы. Это разложение, вообще говоря, не единственно.

Мы будем использовать следующее понятие: G называется P-группой, если функция, тождественно равная единице, может быть аппроксимирована равномерно на каждом компактном подмножестве положительно определенными функциями, равными нулю вне компактных множеств. Если эта аппроксимация может быть осуществлена функциями вида $c * \widetilde{c}(g)$, где $\widetilde{c}(g) = \overline{c(g^{-1})}$ и $c(g)$ равно нулю вне некоторого компактного множества, то мы говорим, что G есть P^1-группа. В частности, можно было бы испробовать функции

$$p(g) = \frac{1}{\nu(C)} I_C * \widetilde{I}_C(g),$$

где $I_C(g)$ — индикатор компактного множества C (см. замечания 5.1.$_1$).

5.2. Анализ Фурье на локально компактных стохастических группах

По аналогии с гл. 3 примем следующее

Определение 5.2. *Преобразование Фурье $\hat{P} = \hat{P}_r$ распределения вероятностей $P \in \mathcal{P}(G)$ на локально компактной группе G определяется соотношением*

$$\hat{P}_r z = \int_{g \in G} U(g) z P(dg),$$

где $r = (\mathcal{H}, U(g))$ — неприводимое представление, а $z \in \mathcal{H}$.

Так как $U(g)z$ сильно непрерывна и ограничена, то нетрудно определить интеграл; его можно понимать как интеграл Бохнера. Заметим, что преобразование Фурье $\hat{P}_r$ есть операторнозначная функция, определенная на множестве всех унитарных неприводимых представлений группы G.

Можно заметить, что вероятностный оператор $Tf(g) = \int_G f(gh)P(dh)$, соответствующий регулярному представлению $f(g) \to R_h f(g) = f(gh)$, обычно не является неприводимым.

Имеет место следующий основной результат.

Теорема 5.2.

а) *Для любого неприводимого унитарного представления* r *преобразование Фурье* $\hat{P}_r$ *является ограниченным линейным оператором в* $\mathcal{H}$.

б) *Если* r *— идентичное представление, то* $\hat{P}_r = I$.

в) *Если известно* $\hat{P}_r$ *для всех неприводимых унитарных представлений* r, *то* P *однозначно определено.*

г) $\widehat{P_1 * P_2} = \hat{P}_1 \cdot \hat{P}_2$.

д) *Пусть* g *имеет распределение* P. *Обозначим через* $\overline{P}$ *распределение вероятностей* g^{-1}, *так что* $\overline{P}(E) = P(E^{-1})$. *Тогда преобразование Фурье распределения* $\overline{P}$ *является сопряженным* $(\hat{P})^*$ *к* $\hat{P}$. *В частности, преобразование Фурье* $\hat{P}$ *является самосопряженным оператором тогда и только тогда, когда* P *есть симметричная мера,* $P = \overline{P}$. *Оно является нормальным оператором тогда и только тогда, когда* $P * \overline{P} = \overline{P} * P$.

е) *Пусть дана последовательность распределений вероятностей* $P_1, P_2, \ldots,$ *слабо сходящаяся к* $P \in \mathcal{P}(G)$. *Тогда имеет место сильная сходимость* $\hat{P}_n \to \hat{P}$. *С другой стороны, из слабой сходимости* $\hat{P}_n \to \hat{P}$ *(с* $P \in \mathcal{P}(G)$*) вытекает слабая сходимость* $P_n \to P$.

Доказательство. Справедливость утверждений а), б) и г) очевидна. Для того чтобы доказать в), предположим, что P_1 и $P_2 \in \mathcal{P}(P)$ и $\hat{P}_1 = \hat{P}_2$ для всех r. Если $p(g)$ — произвольная элементарная положительно определенная функция, то она может быть записана в виде $p(g) = (U(g), z, z)$ с помощью некоторого неприводимого представления $r = (\mathcal{H}, U(g))$. Но тогда при $Q = P_1 - P_2$ имеем

$$\int_G p(g)\, Q(dg) = (\hat{P}_1 z, z) - (\hat{P}_2 z, z) = 0,$$

так что

$$\int_G \varphi(g)\, Q(dg) = 0$$

для любого тригонометрического полинома $\varphi(g)$. Отсюда вытекает (см. разд. 5.1), что $Q = 0$, $P_1 = P_2$.

Для того чтобы доказать д), рассмотрим

$$(\hat{\bar{P}}) = \int_G U(g)\,\bar{P}(dg) = \int_G U(g^{-1})\,P(dg) = \int_G U^*(g)\,P(dg) = (\hat{P})^*.$$

Первую часть утверждения е) проверить нетрудно. Для заданного $x \in \mathscr{H}$ имеем

$$\|(\hat{P}_n - \hat{P})\,x\|^2 = \|\hat{P}_n x\|^2 + \|\hat{P}x\| - 2\,\mathrm{Re}\,(\hat{P}_n x,\ \hat{P}x).$$

Так как подинтегральная функция непрерывна и ограничена, то

$$\lim_{n\to\infty} (\hat{P}_n x,\ \hat{P}x) = \lim_{n\to\infty} \int_G (U(g)\,x,\ \hat{P}x)\,P_n(dg) =$$

$$= \int_G (U(g)\,x,\ \hat{P}x)\,P(dg) = (\hat{P}x,\ \hat{P}x) = \|\hat{P}x\|^2.$$

Так как имеет место слабая сходимость $P_n \to P$, то отсюда вытекает $\bar{P}_n \to \bar{P}$ и слабая сходимость $\bar{P}_n * P_n \to \bar{P} * P$ (см. замечания 2.2.$_2$). Следовательно,

$$\lim_{n\to\infty} \|\hat{P}_n x\| = \lim_{n\to\infty} (\hat{P}_n x,\ \hat{P}_n x) = \lim_{n\to\infty} (\hat{P}_n^* \hat{P}_n x,\ x) =$$

$$= (\hat{P}^* \hat{P}_x,\ x) = (\hat{P}x,\ \hat{P}x) = \|\hat{P}x\|^2,$$

а отсюда вытекает сильная сходимость $(\hat{P}_n - \hat{P})\,x \to 0$.

Вторая часть утверждения е) требует более сложных рассуждений. Так как $P_n\,(G) = 1 = P\,(G)$, слабая сходимость P_n к P вытекает из L-слабой сходимости, т. е. из сходимости в смысле равенства

$$\lim_{n\to\infty} \int_G f(g)\,P_n(dg) = \int_G f(g)\,P(dg)$$

для каждой $f \in L\,(G)$. Пусть сначала $p(g)$ есть произвольная (непрерывная) положительно определенная функция. Как обычно, мы относим ее к некоторому унитарному представлению $\{\mathscr{H},\ U\,(g)\}$: $p\,(g) = (U\,(g)\,x,\ x)$. Разлагая $U\,(g)$ на неприводимые компоненты $U_\delta(g)$, имеем (см. разд. 5.1)

$$p\,(g) = \int_{\delta\in D} (U_\delta(g)\,x_\delta,\ x_\delta)\,\varrho\,(d\delta),$$

и по теореме Фубини и в силу ограниченной сходимости

$$\int_G p(g)\,P_n(dg) = \int_{\delta\in D} (\hat{P}_{n,\,\delta}x_\delta,\ x_\delta)\,\varrho(d\delta) \to$$

$$\to \int_{\delta\in D} (\hat{P}_\delta x_\delta,\ x_\delta)\,\varrho(d\delta) = \int_G p(g)\,P(dg).$$

Остается теперь доказать, что это сохраняет силу, если p заменить произвольной функцией из $L(G)$. Пусть $f \in L$. Тогда f может быть равномерно аппроксимирована на G композициями вида $f * \tilde{\varepsilon}$, где $\varepsilon \in L(G)$ (см. замечания 5.2.$_4$). Но $f * \tilde{\varepsilon}$ может быть линейно выражена через $(f \pm \varepsilon) * \widetilde{(f \pm \varepsilon)}$ и $(f \pm i\varepsilon) * \widetilde{(f \pm i\varepsilon)}$. Последние функции вида $h * \tilde{h}$ являются положительно определенными и непрерывными, что и заканчивает доказательство.

Идемпотенты описываются аналогично случаям компактных и коммутативных групп.

Теорема 5.2.1. *P является идемпотентной вероятностной мерой на G тогда и только тогда, когда она является нормированной мерой Хаара на компактной подгруппе.*

Доказательство. Так как $(\hat{P}_r)^2 = \hat{P}_r$, то для любого $x \in \mathscr{H}_r$ имеем $\hat{P}_r z = z$, если положить $z = \hat{P}_r x$. Но, как и раньше, отсюда вытекает, что $U_r(g)\,z = z$ для любого $g \in s(P)$. В силу произвольности x отсюда вытекает, что $U_r(g)\,\hat{P}_r = \hat{P}_r$ в $s(P)$, так что

$$s(P) \subset \bigcap_r \{g \mid U_r(g)\,\hat{P}_r = \hat{P}_r\}.$$

Множество, стоящее справа, является замкнутой подгруппой группы G. Как обычно, можно предполагать, что G уже выбрана как наименьшая замкнутая подгруппа (исходной группы), содержащая $s(P)$; тогда указанное множество есть просто G. Рассмотрим теперь вероятностную меру Q, определенную соотношением $Q(E) = P(hE)$. Тогда

$$\hat{Q}_r = \int_G U_r(g)\,Q(dg) = \int_G U_r(g)\,P(h\,dg) =$$

$$= \int_G U_r(h^{-1}g)\,P(dg) = U_r(h^{-1})\,\hat{P}_r = \hat{P}_r.$$

В силу теоремы единственности отсюда следует, что $Q(E) = P(hE) = P(E)$; P есть нормированная мера Хаара на G, откуда вытекает, что G компактна. Оставшаяся часть доказательства стандартна.

Имеет место следующая простая лемма.

Лемма 5.2.1. *Пусть* $P \in \mathcal{P}(G)$. *Если существует неприводимое представление* $r = (\mathcal{H}, U(g))$ *(отличное от идентичного представления), такое, что* $\hat{P}_r z = e^{i\theta} z$ *для некоторого нетривиального* $z \in \mathcal{H}$, *то носитель* $s(P)$ *содержится в классе смежности по собственной замкнутой подгруппе группы* G.

Доказательство является повторением доказательства леммы 3.2.1. Нужно только отметить, что условие леммы формулируется не в терминах нормы оператора $\hat{P}_r$, а в терминах собственных значений (преобразования $\hat{P}_r$), по модулю равных единице. Если мы предполагаем только, что $\|\hat{P}\| = 1$, то мы знаем, что существует последовательность векторов $z_1, z_2, \ldots \in \mathcal{H}$, $\|z_n\| = 1$, такая, что $\|\hat{P} z_n\| \to 1$. Обозначая $p_n(g) = (U(g) z_n, z_n)$, имеем

$$\|\hat{P} z_n\|^2 = \left(\int_G U(g) z_n P(dg), \int_G U(g) z_n P(dg) \right) =$$

$$= \int_G \int_G (U(g^{-1}h) z_n, z_n) P(dg) P(dh) =$$

$$= \int_G (U(g) z_n, z_n) Q(dg) = \int_G P_n(g) Q(dg),$$

где $Q = \overline{P} * P$. Так как $p_n(e) = 1$ и $|p_n(g)| \leqslant 1$, то мы можем извлечь подпоследовательность n_ν, такую, что

$$p_{n_\nu}(g) = (U(g) z_{n_\nu}, z_{n_\nu}) \to 1$$

почти всюду относительно меры Q, или, что эквивалентно, имеет место почти всюду сильная сходимость $U(g) z_{n_\nu} - z_{n_\nu} \to 0$. Но если эта сходимость имеет место для g_1 и g_2:

$$U(g_1) z_{n_\nu} - z_{n_\nu} \to 0,$$

$$U(g_2) z_{n_\nu} - z_{n_\nu} \to 0,$$

то

$$U(g_2^{-1})[U(g_1)-U(g_2)]\,z_{n_\nu}=U(g_2^{-1}g_1)\,z_{n_\nu}-z_{n_\nu}\to 0,$$

так что мера Q имеет всю массу сосредоточенной на подгруппе

$$\{g \mid U(g)\,z_{n_\nu}-z_{n_\nu}\to 0\}.$$

К сожалению, это не дает большой информации.

Заметим в скобках, что среди однородных процессов на G можно следующим образом выделить специальный класс, имеющий некоторый интерес.

Теорема 5.2.2. *Пусть P_t есть однородный процесс разрывного типа на локально компактной группе G. Тогда он является сложным пуассоновским процессом на G.*

Доказательство и обозначения аналогичны использованным в разд. 2.3.

Перед тем как перейти к собственно предельным теоремам, сделаем несколько замечаний относительно сходимости по вероятности для стохастических (локально компактных) групп. Пусть γ_n, $n=1, 2, \ldots$ есть последовательность случайных элементов, принимающих значения из G. Предположим, что эта последовательность сходится по вероятности к некоторому случайному элементу γ, т. е. для каждой окрестности N элемента e и любого $\varepsilon>0$ имеем

$$P\{\gamma_n^{-1}\gamma\in N\}>1-\varepsilon$$

для всех n, бо́льших некоторого $n_0=n_0(N, \varepsilon)$. Тогда, конечно, двойная последовательность $\gamma_n^{-1}\gamma_m$ сходится по вероятности к e. Более существенно то, что имеет место и обратное утверждение (критерий Коши).

Теорема 5.2.3. *Если $\gamma_n^{-1}\gamma_m$ сходится по вероятности к e при n и m, стремящихся к бесконечности, то последовательность γ_n сходится по вероятности.*

Доказательство. Мы можем действовать, как в случае действительной прямой, но с необходимыми изменениями. Выберем последовательность положительных чисел $\varepsilon_1, \varepsilon_2, \ldots,$ удовлетворяющую условию $\sum_1^\infty \varepsilon_\nu<\infty$, и убывающую последовательность $N_1, N_2, \ldots$ окрестностей элемента e, такую, что $N_\nu^2 \subset N_{\nu-1}$ и $\bigcap\limits_\nu N_\nu = e$. Выберем натуральные n_ν так, чтобы

$$P\{\gamma_{n_\nu}^{-1}\gamma_{n_{\nu+1}} \in N_\nu\} > 1-\varepsilon_\nu.$$

Рассмотрим множества

$$S_r = \bigcap_{\nu=r}^{\infty} \{\gamma_{n_\nu}^{-1}\gamma_{n_{\nu+1}} \in N_\nu\}$$

и

$$S = \bigcup_{r=1}^{\infty} S_r.$$

Тогда

$$P(S_r^*) \leqslant \sum_{\nu=r}^{\infty} P\{\gamma_{n_\nu}^{-1}\gamma_{n_{\nu+1}} \overline{\in} N_\nu\} \leqslant \sum_{\nu=r}^{\infty} \varepsilon_\nu,$$

так что

$$P(S_r) \geqslant 1 - \sum_{\nu=r}^{\infty} \varepsilon_\nu \uparrow 1$$

при $r \to \infty$, откуда $P(S) = 1$. Но если имеет место событие S, что происходит почти наверное, то $\gamma_{n\nu}$ есть последовательность Коши. Действительно, если заданы N и $\varepsilon > 0$, то можно найти натуральное p такое, что $N_p \subset N$ и (для $\nu < \mu$), используя включение $N_\nu^2 \subset N_{\nu-1}$, получаем

$$\gamma_{n_\nu}^{-1}\gamma_{n_\mu} = \gamma_{n_\nu}^{-1}\gamma_{n_{\nu+1}}\gamma_{n_{\nu+1}}^{-1}\gamma_{n_{\nu+2}}\cdots\gamma_{n_{\mu-1}}^{-1}\gamma_{n_\mu} \in N_\nu N_{\nu+1}\cdots$$
$$\cdots N_{\mu-1} \subset N_{\nu-1} \subset N_p \subset N$$

для всех $\nu, \mu > p$. Каждая локально компактная топологическая группа является полной, так что существует такой элемент γ (см. замечания 5.2.$_1$), что $\gamma_{n_\nu} \to \gamma$. Но так как γ_{n_ν} сходится по вероятности к γ при $n \to \infty$ и так как $\gamma_n^{-1}\gamma = \gamma_n^{-1}\gamma_{n_\nu}\gamma_{n_\nu}^{-1}\gamma$, то из $\gamma_n^{-1}\gamma_{n_\nu} \to e$ вытекает, что γ_n сходится по вероятности к γ.

Заметим, что попутно мы установили существование подпоследовательности, сходящейся почти наверное.

Пусть теперь $g_1, g_2, \ldots$ — независимые случайные элементы группы G. Мы не предполагаем, что они имеют одинаковое распределение. Образуем произведение $\gamma_n = g_1 g_2 \ldots g_n$. Было бы удобно иметь аналитический критерий для сходимости таких произведений.

Теорема 5.2.4. *Произведения $\gamma_n = g_1 g_2 \ldots g_n$ сходятся по вероятности при $n \to \infty$, если для всех неприводимых представлений имеет место неравенство*

$$\sum_{n=1}^{\infty} \| \hat{P}_n - I \| < \infty,$$

где $\hat{P}_n$ — преобразование Фурье для g_n.

Доказательство. Записывая $\hat{P}_n = I + \Delta_n$, рассматриваем преобразование Фурье $Q_{n,m}$ случайного элемента $g_{n+1} g_{n+2} \ldots g_m$ группы

$$Q_{n,m} = \hat{P}_{n+1} \hat{P}_{n+2} \ldots \hat{P}_m = \prod_{n+1}^{m} (I + \Delta_\nu) =$$

$$= I + \sum_{\nu=n+1}^{m} \Delta_\nu + \sum_{n+1 \leqslant \nu < \mu \leqslant m} \Delta_\nu \Delta_\mu + \ldots .$$

Обозначая $d_n = \|\Delta_n\|$, имеем

$$\| Q_{n,m} - I \| \leqslant \sum_{n+1}^{m} d_\nu + \frac{1}{2} \left(\sum_{n+1}^{m} d_\nu \right)^2 + \frac{1}{6} \left(\sum_{n+1}^{m} d_\nu \right)^3 + \ldots =$$

$$= \exp \left(\sum_{n+1}^{m} d_\nu \right) - 1.$$

В силу предположения теоремы $\sum_{n+1}^{m} d_\nu$ стремится к нулю при $n \to \infty$ и $Q_{n,m}$ стремится к I — преобразованию Фурье для δ_e. Так как $\gamma_n^{-1} \gamma_m = g_{n+1} g_{n+2} \ldots g_m$ стремится по вероятности к e при n и m, стремящихся к бесконечности, мы можем применить теорему 5.2.3 и получить утверждение теоремы 5.2.4 (см. замечания 5.2.$_2$).

Еще одно замечание по поводу произведений независимых элементов группы. Рассматриваются различные способы сходимости: слабая сходимость распределений вероятностей, сходимость по вероятности, сходимость почти наверное и т. д. На действительной прямой хорошо известны логические связи между различными типами сходимости, и многие из них лишь с небольшими изменениями переносятся на общие локально компактные и сепарабельные группы. Однако нужно отметить, что это не всегда так. Например, известно (см. замечания 5.2.$_3$), что для ряда $\sum\limits_1^\infty x_n$ независимых случайных величин x_n, принимающих действительные значения, сходимость почти наверное эквивалентна сходимости распределений частичных сумм. Пусть теперь G есть компактная группа, а g_1 — случайный элемент, распределенный соответственно нормированной мере Хаара, и g_2, g_3, . . .— такая последовательность независимых случайных элементов группы, что $\prod\limits_2^\infty g_n$ не сходится почти наверное. Тогда распределение частичных произведений $\prod\limits_1^\infty g_\nu$ есть нормированная мера Хаара, в то время как сами частичные произведения не сходятся почти наверное.

5.3. Предельные теоремы на локально компактных стохастических группах

В этом разделе мы получим некоторые предельные теоремы для композиций на локально компактных группах (см. также начало гл. 3). Мы убеждены, что использованные методы могут быть применены в значительно более общих ситуациях и что они могут привести к большому числу полезных предельных теорем.

Для заданного элемента $g \in G$ унитарный оператор $U(g)$, соответствующий некоторому неприводимому унитарному представлению ($\mathscr{H}$, $U(g)$), может быть записан (см. замечания 5.3.$_1$) в виде

$$U(g) = \int\limits_{-\pi}^{\pi} e^{i\lambda} dE_g(\lambda),$$

где $E_g(\lambda)$ — разложение единицы (см. замечания). Если $U(g)$ задано, то это разложение определяется единственным образом. Вводя самосопряженный оператор

$$H(g)=\int_{-\pi}^{\pi}\lambda dE_g(\lambda)$$

ограниченной нормы, $\|H(g)\|\leqslant\pi$, можно представить $U(g)$ в виде $U(g)=\exp[iH(g)]$. Заметим, что из-за отсутствия коммутативности мы не можем утверждать, что $H(g_1 g_2)=H(g_1)+H(g_2)$. Но из теории возмущений известно, что так как $U(g)$ непрерывно, то $E_g(\lambda)z$ непрерывно зависит от g, за исключением, быть может, скачков $E_g(\lambda)$. Отсюда вытекает, что функция $H(g)$ сильно непрерывна на G. Можно было ввести семейство операторов по-другому:

$$N(\lambda)=\int_G E_g(\lambda)P(dg).$$

Заметим, что $N(\lambda)$ уже не обязательно является разложением единицы, но, во всяком случае, не убывает по λ в смысле обычного частичного упорядочения линейных операторов в гильбертовом пространстве (см. замечания 5.3.$_3$). Оператор $H(P)$, определяемый ниже, можно записать в виде

$$H(P)=\int_{-\pi}^{\pi}\lambda N(d\lambda).$$

Для любого распределения вероятностей P можно образовать самосопряженный оператор $H(P)$, определенный соотношением

$$H(P)z=\int_G H(g)zP(dg),$$

с правой частью, понимаемой, скажем, как интеграл Бохнера (см. замечания 5.3.$_2$), и с нормой, не превосходящей π. Этот оператор играет роль, аналогичную обычному среднему значению на действительной прямой (или на R^k, или на другом линейном векторном пространстве), однако эта аналогия не распространяется слишком далеко из-за отсутствия скалярного умножения и свойства коммутативности. Как бы то ни было, a priori представляется разумным введение понятия среднего значения с помощью следующего предварительного определения, которое будет в дальнейшем изменено. Пусть дана вероятностная мера $P \in \mathcal{P}(G)$ на локально компактной группе G. Говорят, что элемент $\bar{g} \in G$ есть среднее значение меры P на G, если

$$H(\bar{g}) = H(P) = \int_G H(g)\, P(dg)$$

для каждого неприводимого унитарного представления.

Единственность этого определения ясна, так как если $H(\bar{g}_1) = H(P) = H(\bar{g}_2)$, то $U(\bar{g}_1) = U(\bar{g}_2)$ для каждого $(\mathcal{H}, U(g))$, откуда, как известно, следует, что $\bar{g}_1 = \bar{g}_2$.

Существование среднего значения не обеспечено. Стоит заметить, что мы выражали U как $\exp(iH)$ с ограниченным H. Однако существуют ситуации, в которых естественно работать с представлениями с неограниченными H. Тогда, конечно, интеграл, определяющий $H(P)$, требует более осторожного подхода и понадобится в какой-то форме условие интегрируемости на P. Мы вернемся к этому вопросу в следующем разделе. Читателю будет полезно сравнить это с простым случаем $G = R^k$.

Если P симметрично, то $H(P) = 0$, так как $H(g^{-1}) = -H(g)$ и $\bar{g} = e$.

Это понятие среднего значения в действительности не имеет практического интереса, но может быть модифицировано, и мы выскажем теорему 5.3.1 в других, хотя и похожих, терминах. Рассмотрим треугольную последовательность случайных элементов группы

$$\begin{array}{l} g_1 \\ g_{21},\ g_{22} \\ g_{31},\ g_{32},\ g_{33} \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ g_{n1},\ g_{n2},\ g_{n3},\ \ldots,\ g_{nn} \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ , \end{array}$$

где эти случайные элементы независимы и в пределах n-й строки имеют одинаковое распределение P_n. Мы уже рассматривали оператор

$$N_1 = H(P) = \int_{-\pi}^{\pi} \lambda N(d\lambda).$$

Введем также оператор

$$N_2 = \int_{-\pi}^{\pi} \lambda^2 N(d\lambda).$$

Теорема 5.3.1. *Предположим, что*

$$1)\ H(P_n) = \int_G H(g) P_n(dg) = \frac{H(g_0)}{n} + R_n,\ \| R_n \| = o\left(\frac{1}{n}\right)$$

и

$$2)\ \| N_2^{(n)} \| = \| \int_G H^2(g) P_n(dg) \| = o\left(\frac{1}{n}\right).$$

Тогда

$$\lim_{n\to\infty} P_n^{n*} = \delta_{g_0}.$$

Доказательство. Преобразование Фурье для P_n^{n*} имеет вид

$$(\hat{P}_n)^n = \left(\int_G U(g) P_n(dg)\right)^n = \left(\int_G \exp[iH(g)] P_n(dg)\right)^n =$$

$$= (I + iH(P_n) + A_n)^n,$$

где

$$A_n = \int_{-\pi}^{\pi} [e^{i\lambda} - 1 - i\lambda] N(d\lambda) = \int_{-\pi}^{\pi} [\cos\lambda - 1] N(d\lambda) +$$

$$+ i \int_{-\pi}^{\pi} [\sin\lambda - \lambda] N(d\lambda) = A_n' + A_n''.$$

Так как $|\cos\lambda - 1|$ и $|\sin\lambda - \lambda| \leqslant a\lambda^2$ и так как операторы $N(d\lambda)$ неотрицательны, то

$$-aN_2^{(n)} \leqslant A_n' \leqslant aN_2^{(n)},$$

$$-aN_2^{(n)} \leqslant A_n'' \leqslant aN_2^{(n)}.$$

Но отсюда вытекает, что

$$\| A_n' \|,\ \| A_n'' \| \leqslant a \| N_2^{(n)} \| = o\left(\frac{1}{n}\right).$$

Следовательно, $\|A_n\| = o\,(1/n)$ и

$$(\hat{P}_n)^n = \left[I + \frac{iH(g_0)}{n} + o\left(\frac{1}{n}\right) \right]^n,$$

где член $o\,(1/n)$ заменяет оператор с нормой $o\,(1/n)$. При n, стремящемся к бесконечности,

$$(\hat{P}_n)^n \to \exp\,[iH(g_0)] = U(g_0) = \hat{\delta}_{g_0},$$

чем и завершается доказательство.

Рассмотрим теперь также оператор

$$N_3^{(n)} = \int_{-\pi}^{\pi} |\lambda|^3 N(d\lambda).$$

Читатель обратит внимание на то, что эти операторы N_1, N_2 и т. д. играют роль обычных моментов в классической теории.

Теорема 5.3.2. *Предположим, что случайные элементы имеют среднее значение e, а распределения вероятностей удовлетворяют условиям*

$$1)\quad N_2^{(n)} = \int_{-\pi}^{\pi} \lambda^2 N(d\lambda) = \frac{S}{n} + o\left(\frac{1}{n}\right),$$

$$2)\quad N_3^{(n)} = \int_{-\pi}^{\pi} |\lambda|^3 N(d\lambda) = o\left(\frac{1}{n}\right).$$

Если $\pi \in \mathcal{P}(G)$ *есть распределение вероятностей, для которого* $\hat{\pi} = \exp\,(-S/2)$, *то* $P_n^{n*} \to \pi$.

Доказательство. Так же, как и в предыдущем доказательстве,

$$\hat{P}_n = \int_{-\pi}^{\pi} e^{i\lambda} N(d\lambda) = \int_{-\pi}^{\pi} \left[1 + i\lambda - \frac{\lambda_2}{2}\right] N(d\lambda) +$$

$$+ \int_{-\pi}^{\pi} \left[e^{i\lambda} - 1 - i\lambda + \frac{\lambda^2}{2}\right] N(d\lambda) = I - \frac{S}{2n} + o\left(\frac{1}{n}\right),$$

так что

$$(\hat{P}_n)^n \to \exp\left(-\frac{S}{2}\right) = \hat{\pi}.$$

Заметим, что так как S — неотрицательный оператор, то существует такой самосопряженный оператор B, что $S = B^2$. Тогда $\hat{\pi}$ принимает знакомый вид: $\exp(-B^2/2)$.

В классических предельных теоремах распределения $P_{n\nu}$ часто определяются нормированными случайными величинами $b_{n\nu}(x_{n\nu} - a_{n\nu})$. В настоящей весьма общей ситуации *такая нормирующая операция, как умножение на скаляр $b_{n\nu}$, неприменима.* В этом состоит основная трудность, которую нужно преодолеть перед построением окончательной общей теории. Пусть задана последовательность случайных величин $x_1, x_2, x_3, \ldots$ (для конкретности будем считать, что $x_n = g_1 g_2 \ldots g_n$ задана на стохастической группе), определенных на некотором пространстве X. Для нормировки распределений величин x отображаем X на некоторое другое пространство Y с помощью последовательности функций $y_n = y_n(x)$, и предельная теорема будет утверждать, что распределение $y_n(x)$ сходится при $n \to \infty$. Но такая нормировка может быть произведена множеством различных способов, причем не все из них имеют для нас одинаковое значение. Предположим, что мы можем доказать две различные предельные теоремы $\mathcal{L}^1$ и $\mathcal{L}^2$ с помощью двух отображений $y_n^1 = y_n^1(x)$ и $y_n^2 = y_n^2(x)$, причем в первой теореме предельным распределением $y_n^1(x)$ при $n \to \infty$ является D^1, а во второй предельным распределением $y_n^2(x)$ при $n \to \infty$ является D^2. Если распределение D^1 неатомическое, а D^2 имеет по крайней мере один атом, скажем A, то есть основания отдавать предпочтение предельной теореме $\mathcal{L}^1$ перед теоремой $\mathcal{L}^2$. Действительно, в то время как в предельной теореме $\mathcal{L}^2$ некоторая вероятность «собрана» в A, в теореме $\mathcal{L}^1$ она «распределена», так что можно говорить, что теорема $\mathcal{L}^1$ дает более детальную картину, чем $\mathcal{L}^2$.

Другая возможность заключалась бы в том, чтобы считать, что предельная теорема $\mathcal{L}^1$ не хуже другой теоремы $\mathcal{L}^2$, если D^1 абсолютно непрерывно относительно D^2. Это соглашение сделало бы множество предельных теорем частично упорядоченным.

Практический интерес предельной теоремы заключается в получении аппроксимации вида

$$P_n(x_n \in S_n) = P_n(y_n \in S) \cong D(S),$$

где $S_n = y_n^{-1}(S)$ и отображение предполагается взаимно однозначным.

В следующем разделе мы рассмотрим другой вид нормировки, не дающий, вообще говоря, решения поставленной выше задачи, но все же приводящий к интересным предельным теоремам.

5.4. Предельные теоремы на некоторых полных группах

На действительной прямой, в R^k и в банаховом пространстве нормированные суммы

$$\frac{1}{n}x_1 + \frac{1}{n}x_2 + \ldots + \frac{1}{n}x_n \text{ и}$$

$$\frac{1}{\sqrt{n}}y_1 + \frac{1}{\sqrt{n}}y_2 + \ldots + \frac{1}{\sqrt{n}}y_n,$$

$$y_\nu = x_\nu - \mathbf{E}x_\nu,$$

имеют особенно привлекательные предельные свойства. На общей локально компактной группе нет операции, соответствующей умножению на скаляр $c \cdot x$ в линейном векторном пространстве. Это вызывает существенное затруднение, которое вкратце будет обсуждаться далее в этом разделе.

Рассмотрим группы, в которых однозначно определены корни n-й степени, т. е. такие, что для любых $g \in G$ и натурального числа n существует единственный элемент $h \in G$, удовлетворяющий уравнению $h^n = g$, он обозначается $h = g^{1/n}$. (Такие группы иногда называются полными R-группами.) Это — существенное ограничение, хотя иногда можно вложить заданную группу в группу, обладающую этим свойством. В таких группах можно определить рациональные степени g^r (r — рациональное число) и в случае, если $g^r \to e$ при $r \to 0$, можно ввести произвольные степени g^t (t — действительное число). Сделаем следующие предположения.

а) Для любого действительного t и любого $g \in G$ существует элемент $g^t \in G$, зависящий непрерывно от g и t.

б) $g^0 = e, \quad g^1 = g$.

в) $g^{t+s} = g^t g^s$.

Введем теперь понятие *среднего*, отличное от понятия среднего значения из разд. 5.3. Пусть $\{\mathscr{H}, U(g)\}$ — произвольное, неприводимое и унитарное представление группы G. При фиксированном g операторы $V_t = U(g^t)$ образуют непрерывную группу унитарных операторов

$$V_{t+s} = U(g^{t+s}) = U(g^t g^s) = U(g^t)\, U(g^s) = V_t V_s.$$

Известно (см. замечания 5.4.$_1$), что

$$V_t = \exp[itH(g)] = \int_{-\infty}^{\infty} e^{it\lambda} dF_g(\lambda),$$

где $F_g(\lambda)$ — разложение единицы, соответствующее самосопряженному оператору

$$H(g) = \int_{-\infty}^{\infty} \lambda dF_g(\lambda).$$

Операторы $H(g)$ могут быть не определены на всем $\mathscr{H}$, так как они не обязательно ограничены. Мы будем предполагать, что существует множество $\mathscr{D}$, всюду плотное в $\mathscr{H}$ и такое, что $H(g)\, z$ определено для $z \in \mathscr{D}$ и всех $g \in G$, и что

$$\int_G \| H(g)\, z \| P(dg) < \infty.$$

Тогда можно ввести оператор $H(P)$ в $\mathscr{D}$, положив

$$H(P)\, z = \int_G H(g)\, zP(dg).$$

Определение 5.4.1. *Если существует такой элемент* $g \in G$, *что* $H(\bar{g})\, z = H(P)\, z$ *для всех* $z \in \mathscr{D}$, *то* $\bar{g}$ *называется средним распределения* P *на* G.

Если среднее $\bar{g}$ существует, то оно определено однозначно. Действительно, если $H(\bar{g}_1) = H(P) = H(\bar{g}_2)$, то операторы $e^{iH(\bar{g}_1)}$ и $e^{iH(\bar{g}_2)}$ унитарны ($=U(\bar{g}_1)$ и $U(\bar{g}_2)$), определены всюду и равны на $\mathscr{H}$, следовательно, $\bar{g}_1 = \bar{g}_2$.

Если при распределении P вся масса сосредоточена на g_0, то среднее распределения P есть g_0.

Среднее случайного элемента g^t (t — действительная постоянная) группы равно $(\bar{g})^t$, если $\bar{g}$ существует.

Если даны два распределения вероятностей P_1 и P_2 на G такие, что операторы $H(P_1)$ и $H(P_2)$ коммутируют и их средние $\bar{g}_1$ и $\bar{g}_2$ существуют, то среднее распределения $p_1P_1 + p_2P_2$, $0 \leqslant p_1, p_2 \leqslant 1$, $p_1 + p_2 = 1$ равно $(\bar{g}_1)^{p_1} (\bar{g}_2)^{p_2}$. В частности, если распределение P имеет среднее $\bar{g}$, то «приведенное» распределение $\frac{1}{2} P + \frac{1}{2} \delta_{(\bar{g})^{-1}}$ имеет среднее e.

Как было указано выше, операторы $H(g)$ не обязательно ограничены, и это приводит к дополнительным трудностям. В исключительных случаях (например, когда определенная ниже функция $M(\lambda)$ изменяется на конечном интервале) их можно избежать, но вообще мы должны быть готовы к тому, что это вызовет значительную затрату времени и усилий.

Сформулируем несколько простых результатов, которые могут, однако, оказаться полезными. Преобразование Фурье величины $g^{1/n}$ основано на представлениях

$$U(g^{1/n}) = \int_{-\infty}^{\infty} e^{i\frac{\lambda}{n}} F_g(d\lambda) = \int_{-\pi}^{\pi} e^{i\mu} E_g(d\lambda),$$

где

$$E_g(\lambda'') - E_g(\lambda') = \sum_{k=-\infty}^{\infty} [F_g(n\lambda'' + 2\pi kn) - F_g(n\lambda' + 2\pi kn)]$$

при $-\pi \leqslant \lambda' < \lambda'' < \pi$. Оператор $N_1 = H(P_n)$, связанный со случайным элементом $g^{1/n}$, принимает вид

$$N_1 = H(P_n) = \int_{-\pi}^{\pi} \lambda N^{(n)}(d\lambda) = \int_{-\infty}^{\infty} \left[\frac{\lambda}{n}\right] M(d\lambda),$$

где
$$M(\lambda)=\int_G F_g(\lambda)P(dg).$$

Вспомним, что P есть распределение вероятностей элемента g стохастической группы. Символом $[x]$ обозначена периодическая функция с периодом 2π, определенная на интервале $[-\pi, \pi)$ соотношением $[x]=x$. Аналогично имеем
$$N_2=\int_{-\pi}^{\pi}\lambda^2 N^{(n)}(d\lambda)=\int_{-\infty}^{\infty}\left[\frac{\lambda}{n}\right]^2 M(d\lambda).$$

Вспоминая теорему 5.3.1, получаем непосредственно следующий результат.

Теорема 5.4.1. *Предположим, что*

$$1)\ M_{[1]}=\int_{-\infty}^{\infty}\left[\frac{\lambda}{n}\right]M(d\lambda)=\frac{1}{n}\int_{-\infty}^{\infty}\left[\frac{\lambda}{n}\right]F_{g_0}(d\lambda)+R_n,$$
$$\|R_n\|=o\left(\frac{1}{n}\right)$$

и

$$2)\ \|M_{[2]}\|=\left\|\int_{-\infty}^{\infty}\left[\frac{\lambda}{n}\right]^2 M(d\lambda)\right\|=o\left(\frac{1}{n}\right).$$

Тогда имеет место сходимость $g_2^{1/n}g_2^{1/n}\ldots g_n^{1/n}\to g_0$ *по вероятности, или, что эквивалентно,* $P_n^{n*}\to\delta_{g_0}$.

В качестве иллюстрации приведем без доказательства следствие этой теоремы. Пусть P — симметричное распределение вероятностей, для которого

$$\text{а)}\ \left\|\int_{-\eta n}^{\pi n}\lambda^2 M(d\lambda)\right\|=o(n),$$

$$\text{б)}\ \left\|\int_{\pi n}^{\infty}M(d\lambda)\right\|=o\left(\frac{1}{n}\right).$$

Тогда $P_n^{n*}\to\delta_{g_0}$.

Можно получить также не вполне удовлетворительное распространение теоремы 5.3.2.

Теорема 5.4.2. *Предположим, что существует такой положительный оператор* $\hat{\pi}$, *что*

$$1)\ \| M_{[1]} \| = o\left(\frac{1}{n}\right),$$

$$2)\ M_{[2]} = I - (\hat{\pi})^{1/n} + o\left(\frac{1}{n}\right),$$

$$3)\ \| M_{[3]} \| = \left\| \int_{-\infty}^{\infty} \left| \left[\frac{\lambda}{n}\right] \right|^3 M(d\lambda) \right\| = o\left(\frac{1}{n}\right).$$

Тогда имеет место слабая сходимость $P_n^{n*} \to \pi$.

Доказательство. Сравниваем с доказательством теорем 5.3.1 и 5.3.2 и замечаем, что

$$(\hat{P}_n)^n = \left[I + M_{[2]} + o\left(\frac{1}{n}\right)\right]^n \to \hat{\pi}.$$

Теоремам этого раздела нельзя придавать бóльшее значение, чем предварительным и пробным попыткам формулировать теорию. Их недостатки должны быть очевидны читателю и могут побудить его попытаться их устранить. Можно было бы предпочесть формулировать условия в терминах операторов

$$M_p = \int_{-\infty}^{\infty} \left(\frac{\lambda}{n}\right)^p M(d\lambda),$$

вместо того чтобы делать это в терминах операторов $M_{[p]}$. В настоящее время известно только, как это сделать с помощью сложных и громоздких предположений. Требуется больший опыт в обращении с предельными теоремами на специальных группах, для того чтобы могла быть сформулирована соответствующая теория.

До этого можно в качестве ориентира использовать результаты, подобные приведенным выше. Обычная процедура заключается в том, что сначала определяются разложения $E_g(\lambda)$ и затем вычисляются обобщенные разложения $M(\lambda)$ и $N(\lambda)$. Последние используются затем для образования нужных операторов M_1, $M_{[1]}$, N_1 и т. д. Так мы получим указание, какого рода предельные теоремы нужно ожидать.

5.5. Примеры

5.5.1. Для ясности может быть стоит начать с особенно простого случая группы вращений окружности $T^1 = \{\varphi \mid -\pi \leqslant \varphi < \pi\}$. Неприводимые унитарные представления $U(g)$ в этом случае особенно просты — это $e^{i\nu g}$, $\nu = 0, \pm 1, \pm 2, \ldots$ Поэтому имеем:

$$H_0(g) \equiv 0,$$

$$H_1(g) = g,$$

$$H_2(g) = \begin{cases} 2g, & -\dfrac{\pi}{2} \leqslant g < \dfrac{\pi}{2}, \\ 2g - 2\pi, & \dfrac{\pi}{2} \leqslant g < \pi, \\ 2g + 2\pi, & -\pi \leqslant g < -\dfrac{\pi}{2}; \end{cases}$$

$$H_3(g) = \begin{cases} 3g, & -\dfrac{\pi}{3} \leqslant g < \dfrac{\pi}{3}, \\ 3g - 2\pi, & \dfrac{\pi}{3} \leqslant g < \dfrac{2}{3}\pi, \\ 3g - 4\pi, & \dfrac{2}{3}\pi \leqslant g < \pi, \\ 3g + 2\pi, & -\dfrac{2}{3}\pi \leqslant g < -\dfrac{\pi}{3}, \\ 3g + 4\pi, & -\pi \leqslant g < -\dfrac{2}{3}\pi; \end{cases}$$

и т. д.

Для того чтобы проиллюстрировать теорему 5.3.1, мы можем рассмотреть равномерное распределение R_n на интервалах $\left(\dfrac{a}{n}, \dfrac{b}{n}\right)$. Тогда для любого ν и достаточно больших значений n

$$\int_G H_\nu(g)\, R_n(dg) = \frac{\nu}{n} \frac{a+b}{2} = \frac{1}{n} H_\nu\left(\frac{a+b}{2}\right)$$

и

$$\int\limits_G H_\nu^2(g)\,R_n(dg)=\frac{\nu^2}{n^2}\,\frac{b^3-a^3}{3(b-a)}=O\left(\frac{1}{n^2}\right).$$

В этом случае теорема, конечно, применима и $R_n^{n*}\to\delta_{(a+b)/2}$. Беря $b=-a$ (так что $\bar g=e$) и рассматривая большие n, получаем

$$\int\limits_G H_\nu^2(g)\,R_{\sqrt{n}}(dg)=\frac{\nu^2b^2}{3n}$$

и

$$\left\|\int\limits_{-\pi}^{\pi}|\lambda|^3 M(d\lambda)\right\|=\int\limits_G |H_\nu(g)|^3 R_{\sqrt{n}}(dg)=$$
$$=O\left(\frac{1}{n^{3/2}}\right)=o\left(\frac{1}{n}\right),$$

так что можно ожидать, что $R^{n*}_{\sqrt{n}}$ сходится по вероятности к распределению вероятностей π, преобразованием Фурье которого является функция

$$\hat\pi=\exp\left(-\frac{\nu^2b^2}{6}\right).$$

Легко видеть, что если x есть случайная величина, принимающая действительные значения и распределенная нормально со средним 0 и стандартным отклонением σ, то $g\equiv x$, сведенному к интервалу $(-\pi,\pi)$ по модулю 2π, имеет преобразование Фурье

$$\int\limits_{-\pi}^{\pi} e^{i\nu g}P(dg)=\frac{1}{\sqrt{2\pi}\sigma}\int\limits_{-\infty}^{\infty}\exp\left(i\nu g-\frac{x^2}{2\sigma^2}\right)dx=$$
$$=\exp\left(-\frac{\sigma^2\nu^2}{2}\right).$$

В нашем случае $\sigma=b/\sqrt{3}$. Предельным распределением является, конечно, просто нормальное распределение на единичной окружности (см. также 4.4.1).

Остановимся на минуту и подумаем об этом примере. При применении теоремы 5.3.1 мы проверяем выполнение двух условий относительно моментов первого и второго порядков. Но в нашем случае пространство $\mathscr{H}$, в котором действуют наши унитарные представления, является конечномерным, и это дает возможность сформулировать эти условия в более простой форме. В матричной записи

$$U(g) = \{u_{\nu\mu}(g);\ \nu,\ \mu = 1,\ 2,\ \ldots, d\}.$$

Предположим, что

$$\int_G u_{\nu\mu}(g)\, P_n(dg) = \delta_{\nu\mu} + \frac{a_{\nu\mu}}{n} + o\left(\frac{1}{n}\right),\quad A = \{a_{\nu\mu}\}.$$

Тогда, очевидно,

$$\lim_{n\to\infty} (\hat{P}_n)^n = \lim_{n\to\infty}\left[I - \frac{A}{n} + o\left(\frac{1}{n}\right)\right]^n = \exp A.$$

Если $e^A = U(g_0)$, получаем теорему 5.3.1, а если $e^A = \hat{\pi}$, получаем теорему 5.3.2. В частности, введем функцию

$$u(g) = \| U(g) - I \|$$

и потребуем, чтобы

$$\int_G u(g)\, P_n(dg) = o\left(\frac{1}{n}\right).$$

Тогда

$$\| \hat{P}_n - I \| \leqslant \int_G \| U(g) - I \|\, P_n(dg) = o\left(\frac{1}{n}\right),$$

так что $P_n^{n*} \to \delta_e$. Мы могли бы также образовать функцию, построенную по всем представлениям одновременно,

$$v(g) = \sum_1^\infty k_\nu \| U_\nu(g) - I \|,\ k_\nu > 0,\ \sum k_\nu < \infty.$$

Тогда из

$$\int v(g)\, P_n(dg) = o\left(\frac{1}{n}\right)$$

снова вытекает, что $P_n^{n*} \to \delta_e$.

5.5.2. Более интересный, но и более сложный пример дает следующая стохастическая группа, не являющаяся ни компактной, ни коммутативной. Рассмотрим группу линейных преобразований действительной прямой $x \to \alpha x + \beta$, где $\alpha > 0$ (см. замечания 5.5.2).

Элемент этой группы запнсываем в виде $g = (\alpha,\ \beta)$ и определяем окрестности очевидным образом. Группа эта, G, локально компактна. Ее инвариантная справа мера определяется соотношением

$$\nu(dg) = \frac{d\alpha d\beta}{\alpha}.$$

Известно, что ее неприводимые унитарные представления состоят из одномерных представлений вида α^{it}, где t действительно, и некоторых бесконечномерных.

Одномерные представления не будут описываться здесь, так как они ведут себя так же, как и в классическом случае. Пусть H^+ есть гильбертово пространство функций $f(\lambda)$, определенных на положительной части действительной прямой, $0 < \lambda < \infty$, и интегрируемых с квадратом относительно меры Лебега на этой полупрямой; скалярное произведение определяется как обычно. Полагая

$$U^+(g) f(\lambda) = e^{i\lambda\beta} f(\lambda\alpha) \sqrt{\alpha},$$

имеем

$$\| U^+(g) f \|^2 = \int_0^\infty | f(\lambda\alpha) |^2 \alpha d\lambda = \int_0^\infty | f(\mu) |^2 d\mu = \| f \|^2,$$

и $U(g)$ унитарно. Далее, для $g_\nu = (\alpha_\nu, \beta_\nu)$, $\nu = 1,2$, имеем

$$U(g_1) U(g_2) f(\lambda) = U(g_1) e^{i\lambda\beta_2} f(\lambda\alpha_2) \sqrt{\alpha_2} =$$
$$= \exp(i\lambda\beta_1 + i\lambda\alpha_1\beta_2) f(\lambda\alpha_1\alpha_2) \sqrt{\alpha_1\alpha_2} = U(g) f(\lambda),$$

где $g = g_1 g_2$, так что $U(g)$ есть представление. Ясно также, что $U(g)$ непрерывно. Аналогично вводятся H^- и $U^-(g)$. Известно, что описанные представления образуют вместе множество всех неприводимых и унитарных представлений.

Покажем, что G обладает P-свойством. Рассмотрим множество

$$S = \left\{ g \,\middle|\, \frac{1}{A} < \alpha < A, \ -B < \beta < B \right\}, \quad A > 1$$

и его индикатор $I(g) = I_s(g)$. Функция

$$\varphi(g) = \frac{\int\limits_G I(y) I(yg) \nu(dy)}{\nu(s)}$$

является нормированной положительно определенной. Числитель последнего выражения равен

$$\nu\{y \mid y \in S, \ yg \in S\} = 2 \log A (1 + O(1)) 2B (1 + O(1)) =$$
$$= \nu(S)(1 + o(1)),$$

если $A = o(B)$ при A и B, стремящихся к бесконечности. Это означает, что положительно определенная функция $\varphi(g)$ приближает функцию, тождественно равную единице, нужным образом, и утверждение доказано.

Эта группа обладает свойством единственности корней n-й степени. Действительно, для любого действительного t имеем $g^t = (\alpha_t, \beta_t)$, где

$$\alpha_t = \alpha^t,$$

$$\beta_t = \beta \frac{1-\alpha^t}{1-\alpha}$$

при $\alpha \neq 1$ и

$$\alpha_t = 1,$$

$$\beta_t = t\beta$$

при $\alpha = 1$. Исследуя оператор $U(g^t) = \exp(itH(g))$ для малых значений параметра t, находим, что

$$iH(g) f(\lambda) = \left[i\lambda\beta \frac{\log\alpha}{\alpha-1} + \frac{\log\alpha}{2} \right] f(\lambda) + \lambda \log\alpha f'(\lambda)$$

для $f \in \mathscr{D}$, где $\mathscr{D}$ есть множество функций $f(\lambda)$, равных нулю вне конечных интервалов и обладающих непрерывной первой производной.

Рассмотрим такое распределение P на G, что интегралы

$$\int_G \log\alpha P(dg) = a,$$

$$\int_G \beta \frac{\log\alpha}{\alpha-1} P(dg) = b$$

сходятся. В дальнейшем для простоты предполагается, что α и β независимы, хотя это условие не очень существенно. Имеем

$$i \int_G H(g) P(dg) = ibM + \frac{a}{2} I + aMD,$$

где M — оператор умножения на λ, а D — дифференциальный оператор. Тогда

$$i \int_G H(g) P(dg) = iH(\bar{g}),$$

если среднее $\bar{g} = (\xi, \eta)$ выбирается так, что $\xi = e^a$, $\eta = b\frac{e^a - 1}{a}$ при $a \neq 0$ и $\xi = 1$, $\eta = b$ при $a = 0$.

Можно ожидать поэтому, что случайные элементы $\gamma_n = g_1^{1/n} g_2^{1/n} \ldots g_n^{1/n}$ группы сходятся по вероятности к среднему $\bar{g}$. В силу простой структуры группы это можно проверить непосредственно. Полагая $\gamma_n = (\alpha^{(n)}, \beta^{(n)})$, имеем, $\alpha^{(n)} = (\alpha_1 \alpha_2 \ldots \alpha_n)^{1/n}$,

$$\beta^{(n)} = \beta_1 \frac{1 - a_1^{1/n}}{1 - \alpha_1} + \beta_2 \frac{1 - \alpha_2^{1/n}}{1 - \alpha_2} \alpha_1^{1/n} + \ldots + \beta_n \frac{1 - \alpha_n^{1/n}}{1 - \alpha_n} \times \\ \times \alpha_1^{1/n} \alpha_2^{1/n} \ldots \alpha_{n-1}^{1/n}.$$

Логарифмируя, получаем непосредственно из закона больших чисел сходимость по вероятности $\alpha^{(n)} \to \xi = e^a = \exp[\mathbf{E} \log \alpha]$. Полагая

$$b_1^{(n)} = \sum_1^n \beta_\nu \frac{1 - \alpha_\nu^{1/n}}{1 - \alpha_\nu},$$

$$b_2^{(n)} = \frac{1}{n} \sum_1^n \beta_\nu \frac{\log \alpha_\nu}{\alpha_\nu - 1},$$

видим, что имеет место сходимость по вероятности $b_2^{(n)} \to b$. Но

$$b_1^{(n)} - b_2^{(n)} = \sum \beta_\nu \frac{\left[\exp\left(\frac{1}{n} \log \alpha_\nu\right) - 1 - \frac{1}{n} \log \alpha_\nu\right]}{\alpha_\nu - 1},$$

так что из независимости α_ν и β_ν с помощью элементарного, но громоздкого рассуждения получаем, что

$$\mathbf{E} \, | b_1^{(n)} - b_2^{(n)} | \leqslant n \mathbf{E} \, | \beta | \, \mathbf{E} \left| \frac{\exp\left(\frac{1}{n} \log \alpha\right) - 1 - \frac{1}{n} \log \alpha}{\alpha - 1} \right| \to 0$$

и $b_1^{(n)} \to b$ по вероятности. Для того чтобы закончить доказательство, разбиваем сумму, определяющую $\beta^{(n)}$, на такие блоки, чтобы множители вида $(\alpha_1 \alpha_2 \ldots \alpha_\nu)^{1/\nu}$ были постоянны, $\sim \exp\left(\frac{\nu}{n} a\right)$, и применяем приведенное выше рассуждение. Это доказывает сходимость.

Перейдем теперь к распределению элементов $\delta_n = g_1^{1/\sqrt{n}} g_2^{1/\sqrt{n}} \ldots g_n^{1/\sqrt{n}}$ при $a = 0$, $b = 0$, т. е. $\bar{g} = e$. Нужно потребовать существование интегралов

$$\mathbf{E}(\log\alpha)^2 = 1,$$
$$\mathbf{E}\beta^2\left(\frac{\log\alpha}{\alpha-1}\right)^2 = c.$$

Нормировка первого интеграла произвольна и несущественна. Имеем

$$-H^2(g) = \left\{i\lambda\beta\frac{\log\alpha}{\alpha-1} + \frac{\log\alpha}{2} + \lambda\log\alpha D\right\}\left\{i\lambda\beta\frac{\log\alpha}{\alpha-1} + \frac{\log\alpha}{2} + \right.$$
$$\left. + \lambda\log\alpha D\right\} f(\lambda) = A(\alpha, \beta) f(\lambda) + B(\alpha, \beta) f'(\lambda) + C(\alpha, \beta) f''(\lambda),$$

где

$$A(\alpha, \beta) = (\log\alpha)^2\left[-\frac{\lambda^2\beta^2}{(\alpha-1)^2} + 2i\frac{\lambda\beta}{\alpha-1} + \frac{1}{4}\right],$$
$$B(\alpha, \beta) = (\log\alpha)^2\left[\frac{2i\lambda^2\beta}{\alpha-1} + 2\lambda\right],$$
$$C(\alpha, \beta) = \lambda^2(\log\alpha)^2.$$

Следовательно,

$$-M_2 = -\int\limits_G H^2(g)\,P(dg) = A + BD + CD^2,$$

где введены полиномы

$$A = a_2\lambda^2 + a_1\lambda + a_0, \quad B = b_2\lambda^2 + b_1\lambda, \quad C = c_2\lambda^2$$

с коэффициентами

$$a_0 = \frac{1}{4}\mathbf{E}(\log\alpha)^2,$$
$$a_1 = 2i\mathbf{E}(\log\alpha)^2\frac{\beta}{\alpha-1},$$
$$a_2 = -\mathbf{E}(\log\alpha)^2\left(\frac{\beta}{\alpha-1}\right)^2,$$
$$b_1 = 2\mathbf{E}(\log\alpha)^2,$$
$$b_2 = 2i\mathbf{E}(\log\alpha)^2\frac{\beta}{\alpha-1},$$
$$c_2 = \mathbf{E}(\log\alpha)^2.$$

При нашем выборе постоянных получаем:

$$A = -c\lambda^2 + \frac{1}{4}, \quad B = 2\lambda, \quad C = \lambda^2.$$

Для достаточно хорошей функции $f(\lambda)$ введем функцию

$$f(\lambda, t) = \exp\left(-\frac{t}{2} M_2\right) f(\lambda),$$

удовлетворяющую параболическому уравнению

$$\frac{\partial f}{\partial t} = \frac{1}{8} f - \frac{c}{2} \lambda^2 f + \frac{1}{2} (\lambda^2 f')'.$$

Тогда для получения нужной предельной теоремы мы должны выразить оператор $f(\lambda) \to f(\lambda, 1)$ в терминах унитарных представлений. Мы вернемся к этому в гл. 7.

Прежде чем перейти к другому примеру, рассмотрим следующую предельную задачу, отличную от приведенной выше. Рассмотрим элемент $\delta_n = (\alpha^{(n)}, \beta^{(n)}) = (g_1 g_2 \ldots g_n)^{1/n}$. Имеем

$$a^{(n)} = (\alpha_1 \alpha_2 \ldots \alpha_n)^{1/n},$$

$$\beta^{(n)} = [\beta_1 + \beta_2 \alpha_1 + \beta_3 \alpha_1 \alpha_2 + \ldots + \beta_n \alpha_1 \alpha_2 \ldots \alpha_{n-1}] \frac{1 - (\alpha_1 \alpha_2 \ldots \alpha_n)^{1/n}}{1 - \alpha_1 \alpha_2 \ldots \alpha_n}.$$

Первая координата $\alpha^{(n)}$ ведет себя по-прежнему, она сходится по вероятности к среднему геометрическому $\xi = \exp(\mathbf{E} \log \alpha)$, которое предполагается существующим. Для изучения второй координаты рассмотрим сначала случай $\xi < 1$. При этом для последовательности $\beta^{(n)}$ имеет место сходимость распределений (и сходимость по вероятности), и предельным распределением является распределение случайной величины

$$\beta^* = [\beta_1 + \beta_2 \alpha_1 + \beta_3 \alpha_1 \alpha_2 + \ldots] (1 - \xi),$$

где бесконечный ряд сходится почти наверное, если $\mathbf{E}\beta_\nu$ существует, так как с вероятностью единица $(\alpha_1 \alpha_2 \ldots \alpha_n)^{1/n}$ сходится к ξ, так что множители $\alpha_1 \alpha_2 \ldots \alpha_n$ ведут себя асимптотически так же, как ξ^n, $\xi < 1$. С другой стороны, если $\xi > 1$, то для последовательности $\beta^{(n)}$ все же имеет место сходимость распределений (но не сходимость по вероятности) к распределению случайной величины

$$\beta^{**} = \left[\frac{\beta_1}{\alpha_1} + \frac{\beta_2}{\alpha_1 \alpha_2} + \frac{\beta_3}{\alpha_1 \alpha_2 \alpha_3} + \cdots\right] (\xi - 1).$$

Следовательно, в обоих случаях для δ_n имеем невырожденный закон распределения. Сравните это с тем, что имеет место для $\gamma_n = g_1^{1/n} g_2^{1/n} \dots g_n^{1/n}$ и для классического закона больших чисел, когда предельные распределения являются вырожденными.

5.5.3. Рассмотрим теперь спектр оператора T, соответствующего счетной стохастической группе. Пусть G — счетная группа. Введем вероятность p_{gh} перехода от g к h с помощью умножения справа, так что $p_{gh} = P(g^{-1}h)$. Эти вероятности перехода образуют матрицу

$$M = \{p_{h_g};\ g \quad \text{и} \quad h \in G\}.$$

Будет предполагаться, что матрица M симметрична. Ее можно рассматривать как оператор в пространстве l_2 последовательностей f_g, $g \in G$, суммируемых с квадратом.

Ясно, что M идентична оператору T (см. разд. 3.1):

$$Tf_g = \int_G f_{gh} P(dh) = \sum_{h \in G} f_{gh} P(h) = \sum_{k \in G} P(g^{-1}k) f_k =$$
$$= \sum_{k \in G} p_{gk} f_k = (Mf)_g.$$

Спектр M действителен (симметричная мера) и содержится в интервале (—1, 1). Спектральный радиус $r = \max |\lambda|$, где λ принадлежит спектру, имеет специальный интерес и особенно важно знать, меньше ли он единицы или равен единице. (См. также соответствующее обсуждение в разд. 3.1.) Вообще говоря, не легко определить r. Представляет некоторый интерес возможность определения r в следующем случае.

Пусть G — *свободная группа*, порожденная h свободными образующими. Если P имеет равномерное симметричное распределение на этих образующих, то (см. замечания 5.5.3)

$$r = \sqrt{\frac{2h-1}{h^2}}.$$

Эта величина меньше единицы при $h \geqslant 2$. На таких группах нельзя аппроксимировать константу 1 функциями вида $\varphi * \tilde{\varphi}$ способом, описанным ранее в этой главе.

Для конечных или коммутативных счетных групп имеем $r = 1$.

Глава 6

СТОХАСТИЧЕСКИЕ ЛИНЕЙНЫЕ ПРОСТРАНСТВА

6.1. Вероятности на банаховом пространстве

В последней главе мы пытались сформулировать предельные теоремы, аналогичные закону больших чисел и центральной предельной теореме. До некоторой степени это оказалось возможным, но ясно, что можно ожидать более содержательных результатов, если наложить бóльшие требования на структуру группы. В частности, если потребовать, чтобы группа и операция, соответствующая образованию степени, были простыми, то успеха можно достичь значительно легче.

Понятие, которое мы имеем в виду, есть понятие *линейного пространства*, в частности *банахова пространства*. Оно, конечно, коммутативно, и только что упомянутая операция, которая в данном случае является просто умножением на скаляр, дистрибутивна. На первый взгляд можно было бы подумать, что это даст возможность получить желаемые результаты без лишних хлопот простым применением анализа Фурье. Это не так в силу того, что теперь мы имеем дело с пространствами, которые не являются локально компактными. Больше мы не имеем инвариантной меры, а бесконечномерная природа пространства вызывает некоторые не очень приятные последствия, с чем мы столкнемся ниже.

Рассмотрим банахово пространство X (которое обычно считается действительным) с элементами x и нормой $\|x\|$. Двойственное пространство, состоящее из всех непрерывных линейных функционалов $x^*(x)$, обозначается X^*. Мы будем все время предполагать, что и X, и X^* сепарабельны. Читателю следует иметь в виду, что некоторые из результатов этой главы имеют место (или, во всяком случае, известны) только в предположении сепарабельности.

Среди различных банаховых пространств мы будем, конечно, уделять особое внимание сепарабельному *гильбертовому пространству*.

Будем рассматривать вероятностные меры P на X, определенные на σ-алгебре, порожденной открытыми подмножествами пространства X. Тогда любой непрерывный линейный функционал $x^*(x)$ измерим. Для действительного построения меры может оказаться более удобным исходить из конечномерного распределения $x_1^*(x), x_2^*(x), \ldots, x_n^*(x)$, где $x_\nu^* \in X^*$, т. е. использовать алгебру множеств, порожденную цилиндрическими множествами вида

$$\{x \mid (x_1^*(x),\ x_2^*(x),\ \ldots,\ x_n^*(x)) \in B\},$$

где B — произвольное борелевское множество в R^n. Затем мы применяем теорему Колмогорова и распространяем вероятностную меру на соответствующую σ-алгебру $\mathscr{L}$, содержащую все открытые множества (см. замечания 6.1.$_1$). Этот возможный подход основан на предположении, что $x^*(x)$ является P-измеримым для любого $x^* \in X^*$; такой случайный элемент называется L-стохастическим.

Здесь уместно упомянуть о понятии *семейства единственности*. Пусть $\mathscr{U} = \{U\}$ — такое семейство множеств (будем предполагать их борелевскими), что знание значений вероятностной меры $P(U)$ на всех U полностью определяет P. Тогда $\mathscr{U}$ называется семейством единственности. В нашем случае все полупространства $\{x \mid x^*(x) < c\}$ образуют семейство единственности. Для конечномерных евклидовых пространств это давно известно (см. замечания 6.1.$_2$).

Среднее значение m распределения вероятностей P на банаховом пространстве определяется с помощью интеграла Петтиса. Будем говорить, что P имеет среднее значение m, если $x^*(x)$ интегрируема для любого x^* и если существует элемент m, удовлетворяющий соотношению

$$x^*(m) = \int_X x^*(x)\, P(dx)$$

для всех $x^* \in X^*$. Если m существует, то оно единственно, и мы будем обозначать $m = \mathbf{E}x$. Далее ясно, что

1) если $\mathbf{E}x_1$ и $\mathbf{E}x_2$ существуют, то $\mathbf{E}(x_1 + x_2)$ существует и равно сумме $\mathbf{E}x_1$ и $\mathbf{E}x_2$;

2) если x есть постоянный элемент x_0 с вероятностью 1 то $\mathbf{E}x$ существует и равно x_0;

3) если $\mathbf{E}x$ существует, то $\mathbf{E}(cx)$ существует для любой неслучайной константы c и равно $c\mathbf{E}x$;

4) если $\mathbf{E}x$ существует и L есть ограниченный линейный оператор, отображающий X в другое банахово пространство Y, то $\mathbf{E}[L(x)]$ существует и равно $L[\mathbf{E}(x)]$;

5) если $\mathbf{E}\|x\| < \infty$, то $\mathbf{E}x$ существует и

$$\|\mathbf{E}x\| \leqslant \mathbf{E}\|x\|.$$

Эти свойства доказываются так же, как соответствующие свойства интегралов Петтиса (см. замечания $6.1._3$).

Если X — гильбертово пространство, то можно ввести также линейный оператор Q, играющий роль, аналогичную матрице моментов второго порядка для конечномерного евклидова пространства. Если $\mathbf{E}\|x\|^2 < \infty$, то мы определяем Q соотношением

$$(Qz, y) = \mathbf{E}(x z)(x, y),$$

т. е.

$$(Qz, z) = \mathbf{E}(x, z)^2,$$

и аналогично определяем *ковариационный оператор* C соотношением

$$(Cz, z) = \mathbf{E}(x - \mathbf{E}x, z)^2.$$

Оператор Q является эрмитовым, неотрицательным, ограниченным и если e_ν — единичные векторы, направленные по координатным осям, то в силу соотношения $(Qy, z) = \mathbf{E}(x, y)(x, z)$ имеем

$$\sum_{\nu, \mu} [(Qe_\nu, e_\mu)]^2 = \sum_{\nu, \mu} [\mathbf{E}(x, e_\nu)(x, e_\mu)]^2 \leqslant$$

$$\leqslant \Big[\sum_{\nu} \mathbf{E}(x, e_\nu)^2\Big]^2 = (\mathbf{E}\|x\|^2)^2 < \infty,$$

так что Q есть оператор Гильберта — Шмидта и, следовательно, вполне непрерывен. Далее,

$$\sum_{\nu} (Qe_\nu, e_\nu) = \mathbf{E}\|x\|^2 < \infty,$$

так что Q имеет конечный след. То же имеет место и для C. Такие операторы называются S-операторами. Сумме $x_1 + x_2$ двух независимых случайных элементов из X соответствует ковариационный оператор, являющийся суммой ковариационных операторов элементов x_1 и x_2.

Если X представлено в виде пространства l_2 последовательностей, то Q и S — обычные бесконечномерные матрицы моментов.

Для последовательности $x_1, x_2, \ldots$ случайных элементов из X имеем несколько полезных понятий сходимости: слабую сходимость почти наверное $P\{x^*(x_\nu) \to x^*(x)$ для всех $x^* \in X^*\} = 1$ и сильную сходимость почти наверное $P\{\|x_n - x\| \to 0\} = 1$, сходимость по вероятности $P\{\|x_n - x\| > \varepsilon\} \to 0$ для любого положительного ε и, наконец, сходимость в среднем с показателем α, $\mathbf{E}\,\|x_n - x\|^\alpha \to 0$.

Что касается сходимости распределений вероятностей $P_n \to P$, то, пожалуй, как и раньше, наиболее важным понятием является слабая сходимость, означающая, что

$$\int\limits_X f(x)\, P_n(dx) \to \int\limits_X f(x)\, P(dx)$$

для всякой непрерывной и ограниченной функции $f(x)$. Это эквивалентно требованию

$$P_n\{x \mid g(x) \leqslant c\} \to P\{x \mid g(x) \leqslant c\}$$

для любой непрерывной функции $g(x)$ и любого c, для которого правая часть непрерывна. Можно показать, что это имеет место в случае, когда $P_n(U)$ сходится на семействе единственности $\{U\}$, и последовательность $\{P_n\}$ компактна.

Заметим, что $\mathcal{P}(X)$ может быть сделано полным сепарабельным метрическим пространством, в котором определенная ниже метрика соответствует слабой сходимости. Возьмем произвольное замкнутое множество C и образуем открытое множество

$$C^\varepsilon = \{x \mid \inf_{y \in C} \|x - y\| < \varepsilon\}.$$

Определим ε_1 как точную нижнюю грань чисел ε, удовлетворяющих неравенству

$$P_2(C) < P_1(C^\varepsilon) + \varepsilon.$$

Аналогично определяем ε_2 как точную нижнюю грань чисел ε, удовлетворяющих неравенству

$$P_1(C) < P_2(C^\varepsilon) + \varepsilon.$$

Метрика определяется соотношением

$$\varrho(P_1, P_2) = \max(\varepsilon_1, \varepsilon_2)$$

(см. замечания $6.1._4$).

6.2. Анализ Фурье в стохастическом банаховом пространстве

Пусть X есть действительное банахово пространство, а P — некоторое распределение вероятностей, определенное на борелевских множествах пространства X. Естественно следующим образом определить преобразование Фурье, или *характеристический функционал* распределения P.

Определение 6.2. *Характеристическая функция (или функционал) $\hat{P}(x^*)$ есть функция на X^*, определяемая по формуле*

$$\hat{P}(x^*) = \mathbf{E}\exp(ix^*(x)) = \int\limits_{x} \exp(ix^*(x))\, P(dx).$$

Некоторые основные свойства характеристических функционалов устанавливаются легко.

Теорема 6.2.1. *В пространстве $\mathcal{P}(X)$ выполнены следующие свойства.*

1) *Функция $\hat{P}(x^*)$ равномерно сильно непрерывна по x^*, а также слабо непрерывна.*

2) $\hat{P}(0) = 1$.

3) $\widehat{P_1 * P_2}(x^*) = \hat{P}_1(x^*)\cdot\hat{P}_2(x^*)$.

4) *$\hat{P}(x^*)$ есть положительно определенная функция на X^*.*

5) *Если $A_{p+1} = \int\limits_{X} \|x\|^{p+1} P(dx) < \infty$, то*

$$\hat{P}(x^*) = \sum_{\nu=0}^{p} \frac{i^p}{p!}\mathbf{E}(x^*(x))^p + \frac{\theta}{(p+1)!} A_{p+1}, \quad |\theta| \leqslant 1.$$

6) *$\hat{P}(x^*)$ однозначно определяет P.*

7) *Если последовательность P_n слабо сходится к распределению вероятностей P, то $\hat{P}_n(x^*) \to \hat{P}(x^*)$ для всех $x^* \in X^*$.*

Доказательство. Утверждения 2), 3), 4) и 5) доказываются точно так же, как и в случае действительной прямой, а 7) непосредственно следует из определения слабой сходимости.

Для доказательства 1) воспользуемся неравенством

$$|\hat{P}(x_1^*) - \hat{P}(x_2^*)| \leqslant \mathbf{E}\,|\exp(ix_1^*(x)) - \exp(ix_2^*(x))| = \\ = \mathbf{E}\,|1 - \exp(iz^*(x))|,$$

где $z^* = x_1^* - x_2^*$. Выберем множество S, удовлетворяющее неравенству $P(S) > 1 - \delta$ и обладающее ограниченными элементами, $\sup\limits_{x \in S} \|x\| = M < \infty$.

Если выбрать z^* столь близким к нулю, что

$$|1 - \exp(iz^*(x))| \leqslant \delta$$

для всех $x \in S$, то

$$|\hat{P}(x_1^*) - \hat{P}(x_2^*)| \leqslant \int\limits_S |1 - \exp(iz^*(x))|\, P(dx) + 2\delta \leqslant 3\delta.$$

Это доказывает первую часть утверждения 1). Вторая часть проверяется аналогично, так как из слабой сходимости $x_2^* \to x_1^*$ в X^* вытекает $x_2^*(x) - x_1^*(x) \to 0$ для любого $x \in X$ и применима теорема Лебега о сходимости.

Утверждение 6) имеет место в силу того, что масса, распределенная на любом множестве вида $U = \{x \mid x^*(x) \leqslant c\}$, может быть вычислена по $\hat{P}(x^*)$, как и в случае действительной прямой. Но множества U образуют семейство единственности и, следовательно, P однозначно определено.

Прежде чем двигаться дальше, отметим тот простой факт, что нам не нужно искать идемпотентные меры в линейном пространстве, так как таких нет (за исключением тривиальной δ_0). Действительно, если $P \in \mathscr{P}(X)$ — нетривиальная идемпотентная мера, то должно существовать такое $x^* \in X^*$, что $P\{x^*(x) \neq 0\} > 0$. Рассмотрим преобразование Фурье

$$\hat{P}(tx^*) = \mathbf{E}\exp(tx^*(x))$$

для действительных значений t. Так как P идемпотентна, то $(\hat{P})^2 = \hat{P}$, так что $\hat{P}(tx^*)$ может принимать только значения 0 и 1. Но это непрерывная функция переменной t, равная единице при $t = 0$. Следовательно, характеристическая функция случайной величины $x^*(x)$ тождественно равна единице, так что в противоречии с предположением, $x^*(x) \equiv 0$.

Теперь мы встречаемся с существенным затруднением. Нам, хотелось бы, конечно, утверждать, что из $\hat{P}_n(x^*) \to \hat{P}(x^*)$ вытекает слабая сходимость $P_n \to P$. На следующем примере можно убедиться, что, вообще говоря, это не так. Пусть $x_1, x_2, x_3, \ldots$ — последовательность элементов из X, слабо (но не сильно) сходящаяся к элементу x_0. Пусть P_n приписывает меру 1 элементу x_n. Тогда $x^*(x_n) \to x^*(x_0)$ и $\exp[ix^*(x_n)] = \hat{P}_n(x^*) \to \exp[ix^*(x_0)] = \hat{P}_0(x^*)$ для всех $x^* \in X^*$. С другой стороны, пусть $f(x)$ — непрерывная и ограниченная функция, равная 1 в $x = x_0$ и нулю вне шара $S_\varepsilon = \{x \mid \|x - x_0\| < \varepsilon\}$, так что для малого ε найдутся сколь угодно большие n, для которых

$$\int_X f(x)\, P_n(dx) = 0,$$

и, следовательно, $\int_X f(x)\, P_n(dx)$ не сходится к $\int_X f(x) P_0(dx) = 1$. Если X — гильбертово пространство l_2, то можно выбрать x_n как единичный вектор, направленный вдоль n-й координатной оси; это ярко иллюстрирует трудности, вызванные бесконечномерностью пространства. Чтобы преодолеть это, заметим следующее (см. замечания 6.2.₁).

Каждое $P \in \mathcal{P}(X)$ является *компактной мерой*: для любого $\varepsilon > 0$ существует такое компактное множество $C \subset X$, что $P(C) \geqslant 1 - \varepsilon$. Действительно, выберем последовательность $x_1, x_2, x_3, \ldots$, плотную в X, и образуем шары $S_n^k = \{x \mid \|x - x_n\| \leqslant 1/k\}$ с центрами в точках x_n и радиусами $1/k$. Выберем m_1 столь большим, что

$$P\{\bigcup_{n=1}^{m_1} S_n^1\} > 1 - \varepsilon.$$

Затем выберем m_2 столь большим, что

$$P\{(\bigcup_{n=1}^{m_1} S_n^1)\cap(\bigcup_{n=1}^{m_2} S_n^2)\} > 1-\varepsilon,$$

и т. д. Множество

$$C=\bigcap_{i=1}^{\infty}(\bigcup_{n=1}^{m_i} S_n^i)$$

компактно и имеет меру $P(C) \geqslant 1-\varepsilon$.

Введем теперь

Определение 6.2.2. *Вероятностные меры $P_n \in \mathcal{P}(S)$ называются равномерно компактными, если для любого положительного ε существует такое компактное множество $C \subset X$, что $P_n(C) \geqslant 1-\varepsilon$ для всех n.*

Тогда справедлива

Теорема 6.2.2. *Для того чтобы имела место слабая сходимость $P_n \to P$, необходимо и достаточно, чтобы меры P_n были равномерно компактны и чтобы $\hat{P}_n(x^*) \to \hat{P}(x^*)$ для всех $x^* \in X^*$.*

Доказательство предоставляется читателю. Само собой разумеется, что выполнение предположения равномерной компактности в теореме 6.2.2 может оказаться трудно проверяемым, что ограничивает полезность этой теоремы. Для гильбертова пространства $X = l_2$ достаточным является следующее условие (см. замечания 6.2.$_2$). Пусть $x = (x_1, x_2, \ldots)$. Введем величину

$$R_N=\sup_n \mathbf{E}\left[\sum_N^{\infty} x_\nu^2\right].$$

Заметим, что верхняя грань берется по P_n.

Теорема 6.2.3. *Если $P_n \in \mathcal{P}(l_2)$ удовлетворяет условиям*

1) $\lim R_N = 0$,

2) $\hat{P}_n(x^*) \to \hat{P}(x^*)$ *для всех $x^* \in X^*$, то имеет место слабая сходимость $P_n \to P$.*

Доказательство. Рассмотрим компактное множество

$$C = \bigcap_{k=1}^{\infty} E_k,$$

где

$$E_k = \left\{ x \,\middle|\, \sum_{N_k}^{\infty} x_\nu^2 \leqslant \frac{1}{l_k} \right\}.$$

Здесь выбрана возрастающая последовательность натуральных чисел $N_k \uparrow +\infty$ с $N_1 = 1$ и последовательность положительных чисел $l_k \to \infty$ так, что

$$\sum_1^{\infty} l_k R_{N_k} < \varepsilon.$$

Но тогда

$$1 - P_n(C) = P_n(C^*) \leqslant \sum_{k=1}^{\infty} P_n(E_k^*) \leqslant \sum_{k=1}^{\infty} l_k R_{N_k} < \varepsilon,$$

и применима теорема 6.2.2.

Следующая трудность возникает при попытке обращения утверждения 4) теоремы 6.2.1. Это обращение не имеет места без дальнейших ограничений. Действительно, рассмотрим функцию

$$\varphi(x^*) = \exp\left[-\frac{1}{2} \| x^* \|^2 \right], \quad x^* \in l_2.$$

Эта функция непрерывна в сильной топологии и положительно определенная (как и нормальные характеристические функции в конечномерных пространствах). Если φ соответствует мере P на l_2, то распределения координат $x_\nu = (x, e_\nu)$ должны быть независимыми случайными величинами $N(0, 1)$, распределенными нормально. Случайные величины

$$R_n^2 = x_1^2 + x_2^2 + \ldots + x_n^2$$

очевидно сходятся по вероятности к $+\infty$ при $n \to \infty$. Но тогда $P\{\| x \| \leqslant R\} = 0$ для любого R, что невозможно. Можно было бы предположить, что неприятности вызваны использованной формой непрерывности. Замена сильной сходимости слабой не приводит к успеху, что можно показать, модифицируя приведенный выше пример.

Подходящей формой сходимости является сходимость, соответствующая окрестностям

$$(Sx^*, x^*) < 1$$

для произвольных S-операторов. Будем называть это S-топологией (см. замечания 6.2.$_3$).

Теорема 6.2.4. *Пусть $\varphi(x^*)$ — комплекснозначная функция переменного x^* на гильбертовом пространстве X^*, удовлетворяющая условию $\varphi(0) = 1$. Для того чтобы $\varphi(x^*)$ была преобразованием Фурье $\hat{P}(x^*)$ некоторого $P \in \mathscr{P}(X)$, необходимо и достаточно, чтобы φ была положительно определенной и непрерывной в точке $x^* = 0$ в S-топологии.*

Доказательство необходимости. Для $P \in \mathscr{P}(X)$ выберем компактное множество $C \subset X$ такое, что $P(C) > 1 - \varepsilon$. Тогда

$$| 1 - \hat{P}(x^*) | \leqslant 2\varepsilon + \int\limits_C [1 - \cos(x^*, x)] P(dx) + \\ + \int\limits_C |\sin(x^*, x)| P(dx) \leqslant 2\varepsilon + \frac{1}{2} \int\limits_C (x^*, x)^2 P(dx) + \\ + \sqrt{\int\limits_C (x^*, x)^2 P(dx)}.$$

Введем S-операторы соотношением

$$(S_\varepsilon x^*, x^*) = \frac{1}{\varepsilon} \int\limits_C (x^*, x)^2 P(dx)$$

и получим требуемое утверждение.

Доказательство достаточности. Рассмотрим пространство Y всех последовательностей $y = (y_1, y_2, \ldots)$. Знание $\varphi(x^*)$ определяет распределение вероятностей P_n любого конечномерного вектора $(y_1, y_2, \ldots, y_n)$, и, как обычно, мы можем распространить эту вероятностную меру на порождаемую σ-алгебру в Y. Теперь для любого положительного ε можно найти такой оператор S (записанный в диагональной форме в l_2),

$$(Sx^*,\ x^*)=\sum_1^\infty a_\nu (x_\nu^*)^2,\quad \sum_1^\infty a_\nu<\infty,\ a_\nu\geqslant 0,$$

что $|1-\varphi(x^*)|\leqslant\varepsilon$ для x^*, удовлетворяющих неравенству $(Sx^*,\ x^*)<1$. Тогда

$$P\left(\sum_1^n y_\nu^2>1\right)\leqslant\frac{\sqrt{e}}{\sqrt{e}-1}\int_{R^n}\left(1-\exp\left[-\frac{1}{2}\sum_1^n y_\nu^2\right]\right)P_n(dy),$$

поскольку

$$\frac{\sqrt{e}}{\sqrt{e}-1}\left(1-\exp\left(-\frac{1}{2}u\right)\right)\geqslant 1$$

при $|u|>1$. Следовательно,

$$P\left(\sum_1^n y_\nu^2\geqslant 1\right)\leqslant\frac{\sqrt{e}}{\sqrt{e}-1}\times$$

$$\times\frac{1}{(2\pi)^{n/2}}\int_{R^n}\{1-\varphi(y^*)\}\exp\left[-\frac{1}{2}\sum_1^n y_\nu^{*2}\right]dy^*$$

(используются свойства нормальной плотности в R^n и ее характеристической функции) и

$$P\left(\sum_1^n y_\nu^2\geqslant 1\right)\leqslant\frac{\sqrt{e}}{\sqrt{e}-1}\,\varepsilon+\frac{\sqrt{e}}{\sqrt{e}-1}\,\frac{1}{(2\pi)^{n/2}}\times$$

$$\times\int\limits_{\sum_1^n a_\nu (y_\nu^*)^2>1} 2\exp\left[-\frac{1}{2}\sum_1^n y_n^{*2}\right]dy^*=\frac{\sqrt{e}}{\sqrt{e}-1}\left(\varepsilon+2\sum_1^n a_\nu\right),$$

так как

$$\frac{1}{(2\pi)^{n/2}}\int\limits_{\sum_1^n a_\nu (y_\nu^*)>1}\exp\left[-\frac{1}{2}\sum_1^n (y_\nu^*)^2\right]dy^*\leqslant$$

$$\leqslant\frac{1}{(2\pi)^{n/2}}\int_{R^n}\left[\sum_1^n a_\nu (y_\nu^*)^2\right]\exp\left[-\frac{1}{2}\sum_1^n (y_\nu^*)^2\right]dy^*=\sum_1^n a_\nu.$$

Аналогично

$$P\left(\sum_1^n y_\nu^2 \geqslant C\right) \leqslant \frac{\sqrt{e}}{\sqrt{e}-1}\left(\varepsilon + 2\,\frac{\sum_1^n a_\nu}{C}\right).$$

Но отсюда вытекает, что вероятностная мера P, введенная на Y, имеет всю массу распределенной на последовательностях из l_2, $P\,\{y \mid \sum_1^\infty y_\nu^2 < \infty\} = 1$. Можно убедиться в том, что упомянутая выше σ-алгебра содержит все открытые множества пространства l_2 (сравните с рассуждением в замечаниях $6.1._1$). Наконец, преобразование Фурье меры P в l_2 совпадает с $\varphi(x^*)$ для всех $x^* \in X^*$. Это завершает доказательство (см. замечания $6.2._3$).

В других банаховых пространствах положение, конечно, сложнее.

Поучительно вернуться на время к случаю действительной прямой. В качестве исходной точки возьмем следующее элементарное

Предложение. *Пусть* $\varphi(t)$, $-\infty < t < \infty$, *есть измеримая и положительно определенная функция. Если*

$$\varphi(0) = \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} \varphi(t)\,dt,$$

то $\varphi(t)$ *непрерывна.*

Доказательство. Хорошо известно, что $\varphi(t)$ совпадает почти всюду с непрерывной положительно определенной функцией $\varphi_c(t)$, т. е. $\varphi(t) = \varphi_c(t) + \varphi_0(t)$, где $\varphi_0(t) = 0$ почти всюду. Для произвольных t_ν, c_ν введем функции

$$\Phi(t) = \sum_{\nu,\,\mu} c_\nu \bar{c}_\mu \varphi(t + t_\nu - t_\mu),$$

$$\Phi_c(t) = \sum_{\nu,\,\mu} c_\nu \bar{c}_\mu \varphi_c(t + t_\nu - t_\mu).$$

Функция $\Phi(t)$ является положительно определенной, так как для любых s_k, a_k имеем

$$\sum_{k,\,l} a_k \bar{a}_l \Phi(s_k - s_l) = \sum_{k,\,l,\,\nu,\,\mu} a_k c_\nu \overline{a_l c_\mu} \varphi(s_k + t_\nu - s_l - t_\mu),$$

что неотрицательно по определению. Отсюда вытекает неравенство $|\Phi(t)| \leqslant \Phi(0)$, так что

$$\Phi(0) \geqslant \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} \Phi(t)\,dt = \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} \Phi_0(t)\,dt + \Phi_c(0),$$

т. е.

$$\Phi(0) - \Phi_c(0) = \sum_{\nu,\,\mu} c_\nu \bar{c}_\mu \varphi_0(t_\nu - t_\mu) \geqslant 0,$$

и $\varphi_0(t)$ — положительно определенная. Но мы предположили, что

$$\varphi_0(0) = \varphi(0) - \varphi_c(0) = \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} \varphi(t)\,dt - \varphi_c(0) =$$

$$= \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} \varphi_c(t)\,dt - \varphi_c(0) = 0.$$

Следовательно, $\varphi_0(0) = 0$, откуда вытекает, что $\varphi_0(t) \equiv 0$.

Ясно, что наше предложение имеет место и для других сингулярных интегральных ядер.

Вернемся теперь к банаховым пространствам и рассмотрим простой пример. Пусть $X = l_1$ состоит из суммируемых последовательностей $x = (x_1, x_2, \ldots)$, $\|x\| = \sum_1^\infty |x_\nu| < \infty$, а двойственное пространство $X^* = m$ состоит из ограниченных последовательностей $t = (t_1, t_2, \ldots)$, $\|t\| = \sup |t_\nu|$. Пусть случайные величины x_ν независимы и

$$P(x_\nu = 0) = \frac{1}{2},$$

$$P\left(x_\nu = \frac{1}{\nu}\right) = P\left(x_\nu = -\frac{1}{\nu}\right) = \frac{1}{4}.$$

Это не определяет «истинного» распределения в l_1, так как

$$\sum_1^\infty \mathbf{E}\,|x_\nu| = \sum_1^\infty \frac{1}{2\nu} = +\infty,$$

так что $\sum\limits_1^\infty |x_\nu|$ не сходится почти наверное (заметим, что $|x_\nu|$ равномерно ограничены).

С другой стороны, «преобразование Фурье»

$$\varphi(t) = \varphi(t_1, t_2, \ldots) = \mathbf{E}\exp\left(i\sum t_\nu x_\nu\right) = \prod_{\nu=1}^\infty \left[\frac{1}{2} + \frac{1}{2}\cos\frac{t_\nu}{\nu}\right]$$

сходится при любом $t \in m$. Если $\|t\| \to 0$, то $\varphi(t) \to 1$. Это показывает, что сильная непрерывность положительно определенной функции $\varphi(t)$ не обеспечивает того, что $\varphi(t)$ является преобразованием Фурье в том строгом смысле, в каком это понималось.

Теперь мы можем сформулировать условие непрерывности. Пусть Λ — направленное множество вероятностных мер λ в X^*, сходящихся к δ_0 в смысле, точно определяемом ниже. Будем говорить, что $\varphi(t)$ является λ-непрерывной, если

$$\lim_\Lambda \int_{X^*} \varphi(t)\,\lambda(dt) = \varphi(0).$$

Теорема. *Пусть $\varphi(t)$ есть положительно определенная функция на $X^* = m$, удовлетворяющая условию $\varphi(0) = 1$. Она является преобразованием Фурье некоторой вероятностной меры на $X = l_1$ тогда и только тогда, когда она λ-непрерывна для λ-мер вида*

$$\lambda(dt) = \prod_1^n \frac{1}{\pi\varepsilon_\nu\left[1 + \frac{t_\nu^2}{\varepsilon_\nu^2}\right]}\,dt_1 \ldots dt_n\delta_0(dt_{n+1})\ldots,$$

$$\max_{1 \leq \nu \leq n} |\varepsilon_\nu| \to 0.$$

Доказательство. Из предположения λ-непрерывности следует, что $\varphi(t_1, t_2, \ldots, t_n, 0, 0, \ldots)$ непрерывна в R^n (см. доказательство предложения выше), так что все конечномерные распределения P_n величин $x_1, x_2, \ldots$

$\ldots, x_n$ вполне определены. Остается доказать, что $P\{x \mid \sum_1^\infty |x_\nu| < \infty\} = 1$. Введем функцию

$$q(n, \varepsilon) = \int_m \varphi(t)\,\lambda(dt) = \int_{x \in R^n} \int_{t \in R^n} \exp\left[i \sum_1^n t_\nu x_\nu\right] \times$$

$$\times P_n(dx) \prod_1^n \frac{1}{\pi\varepsilon_n\left[1 + \frac{t_\nu^2}{\varepsilon_\nu^2}\right]} dt_1 \ldots dt_n =$$

$$= \int_{x \in R^n} \exp\left[-\varepsilon_n \sum_1^n |x_\nu|\right] P_n(dx) = \mathbf{E} \exp\left[-\varepsilon_n \sum_1^n |x_\nu|\right].$$

Тогда из λ-непрерывности вытекает, что

$$\mathbf{E} \exp\left[-\varepsilon_n \sum_1^n |x_\nu|\right] = q(n, \varepsilon_n) \to 1, \qquad n \to \infty,$$

так что

$$\varepsilon_n \sum_1^n |x_\nu| \to 0 \text{ по вероятности при } n \to \infty.$$

Но отсюда вытекает, что $S_n = \sum_1^n |x_\nu|$ сходится (к конечному пределу) по вероятности, так как иначе для некоторого $\delta_0 > 0$ можно было бы указать такую последовательность $A_n \uparrow +\infty$, что $P(S_n > A_n) > \delta_0$. При выборе $\varepsilon_n = 1/A_n \downarrow 0$ это означает, что $P(\varepsilon_n S_n > 1) > \delta_0$ для больших n, что невозможно. Это доказывает, что $P\{x \mid \sum_1^\infty |x_\nu| < \infty\} = 1$.

Предположим теперь, что P есть «настоящая» вероятностная мера на l_1. Тогда

$$1 \geqslant \int \varphi(t)\,\lambda(dt) = \mathbf{E} \exp\left[-\sum_1^n \varepsilon_\nu |x_\nu|\right] \geqslant$$

$$\geqslant \mathbf{E} \exp\left[-\max \varepsilon_\nu \|x\|\right] \to 1$$

при $n \to \infty$, так как $\|x\|$ есть обычная случайная величина. Следовательно, $\varphi(t)$ является λ-непрерывной.

Это рассуждение имеет общий характер. В случае гильбертова пространства мы можем поступать аналогичным образом, используя в качестве Λ множество независимых и нормальных распределений $N(0, \sigma)$ для t_ν с условием $\sum_1^\infty \sigma_\nu^2 < \infty$. Это привело бы нас опять к S-топологии. В общем случае надо было бы вложить пространство X в бо́льшее пространство X', возможно заданное в координатной форме. Можно было бы тогда попытаться аппроксимировать $I_X(x')$, индикатор множества X в X', более регулярными функциями $I_X^{(n)}$ и выразить $\mathbf{E}I_X(x') = P(X)$ через преобразование Фурье функции $I_X^{(n)}$ и $\varphi(x^*)$.

Рассмотрения этого типа имеют некоторый интерес для выяснения логических трудностей, но окончательные критерии не представляют большой практический интерес из-за их косвенной формы.

Вспомним теперь замечания, сделанные в конце разд. 4.2, и посмотрим, как они могут быть применены в случае гильбертова пространства. Предположим, что наше гильбертово пространство X может быть представлено в виде прямой суммы

$$X = X_1 \oplus X_2 \oplus X_3 \oplus \ldots$$

элементов $x = (x_1, x_2, \ldots)$, где x_k принадлежит гильбертову пространству X_k со скалярным произведением $(a, b)_k$ и нормой $\|a\|_k = \sqrt{(a, a)_k}$. В пространстве X имеем скалярное произведение

$$(x, y) = \sum_{k=1}^{\infty} (x_k, y_k)_k.$$

Предположим, что на подпространствах $Y_n = X_1 \oplus X_2 \oplus \ldots \oplus X_n$ имеется некоторая вероятностная структура и мы хотим ввести ее на X. Положим $\| y_n \|_n^2 = \sum_{\nu=1}^{\infty} \| x_\nu \|_\nu^2$.

Теорема 6.2.5. *Предположим, что на каждом Y_n, $n = 1, 2, \ldots$, согласованным образом определена последовательность вероятностных мер $P_1^{(n)}, P_2^{(n)}, \ldots$. Введем* *ограниченную неубывающую функцию*

$$F_k(u)=\lim_{n\to\infty}P_k^{(n)}\{y_n \mid \|y_n\|_n\leqslant u\}.$$

Если $F_k(+\infty)=\lim\limits_{u\to+\infty}F_k(u)=1$, *то* $P_k^{(n)}$ *определяют распределения вероятностей* P_k *на* X. *Если* $F_k(u)\to 1$ *для любого* $u>0$ *при* $k\to\infty$, *то* $P_k\to\delta_0$.

Если $P_t^{(n)}$ — *однородные процессы, согласованно определенные на* Y_n, *и*

$$F_t(u)=\lim_{u\to\infty}P_t^{(n)}\{y_n \mid \|y_n\|_n\leqslant u\}$$

удовлетворяет условиям $F_t(+\infty)=1$ *и* $F_t(u)\to 1$ *для всех* $u>0$ *при* $t\downarrow 0$, *то* $P_t^{(n)}$ *определяют непрерывный однородный процесс* P_t *на* X.

Доказательство. Мы предположили, что меры P_k^n согласованы, так что $P_t^{(n')}$ есть проекция меры $P_k^{(n'')}$ на $Y_{n'}$ при $n'<n''$. Тогда мы получаем вероятностную меру на пространстве последовательностей $(x_1, x_2, \ldots)$. Рассмотрим функции

$$P_k^{(n+1)}\{y_n \mid \|y_n\|_{n+1}\leqslant u\}=P_k^{(n+1)}\{y_n \mid \|y_n\|_n^2+\|x_{n+1}\|^2\leqslant u\}\leqslant$$
$$\leqslant P_k^{(n)}\{y_n \mid \|y_n\|_n\leqslant u\},$$

образующие невозрастающую последовательность при возрастании n. Следовательно, $F_k(u)$ вполне определена. Надо только показать, что

$$P_k\Big\{x \;\Big|\; \sum_1^\infty\|x_k\|_k^2<\infty\Big\}=1,$$

что следует из

$$P_k\Big\{x \;\Big|\; \sum_1^\infty\|x_k\|_k^2<\infty\Big\}=\lim_{n\to+\infty}P_k\Big\{x \;\Big|\; \sum_1^\infty\|x_k\|_k^2<n\Big\}=$$
$$=\lim_{n\to+\infty}F_k(n)=1.$$

Далее,

$$P_k\{x \mid \|x\|<\varepsilon\}=F_k(\varepsilon)\longrightarrow 1$$

для всех $\varepsilon>0$ при $k\to\infty$, так что $P_k\to\delta_0$.

Оставшаяся часть доказательства завершается применением этого к $P_t^{(n)}$.

Замечание. Для первой части теоремы достаточно (но не необходимо) предположить, что

$$\sigma_k^2 = \lim_{n\to\infty} \int_{Y_n} \| y_n \|_n^2 P_k^{(n)} (dy_n)$$

конечно и $\sigma_k^2 \to 0$ при $k \to \infty$. Аналогично дело обстоит и для второй части.

Интересно применить это к $X = l_2$, где мы выбираем X_n как одномерные пространства, соответствующие координатам в l_2. Для каждого X_n известен вид непрерывного однородного процесса (или, точнее, его инфинитезимального порождающего оператора), и это может быть использовано для построения P_t в $X = l_2$.

6.3. Нормальные распределения в гильбертовом пространстве

Теперь нетрудно видеть, каким должно было бы быть корректное определение *нормального распределения вероятностей* в гильбертовом пространстве X.

Определение 6.3.1. *Нормальное распределение P в гильбертовом пространстве X понимается как распределение, однозначно определенное преобразованием Фурье*

$$\hat{P}(x^*) = \exp\left\{ i(x^*, m) - \frac{1}{2}(Sx^*, x^*) \right\}.$$

Здесь m — фиксированный элемент из X, а S есть S-оператор.

Теорема 6.2.4 показывает, что такое распределение существует, а единственность ясна из теоремы 6.2.1 (утверждение 6). Можно также убедиться, что ограничение на оператор S необходимо.

Теперь мы можем выразить m и S в терминах моментов распределения.

Теорема 6.3.1. *Для нормального распределения элемент m является средним значением, а S — ковариационным оператором.*

Теорема 6.3.2. *Если x — случайный элемент в гильбертовом пространстве X с нормальным распределением, то он может быть записан в виде*

$$x = m + \sum_{\nu} \xi_\nu e_\nu,$$

где e_ν — ортогональные единичные векторы в X, а ξ_ν — независимые случайные величины с распределениями $N(0, \sigma_\nu)$, где σ_ν^2 — собственные значения оператора S. Эта сумма конечна или счетна и сходится (сильно) с вероятностью единица.

Доказательство теоремы 6.3.2. Выберем ортонормальные векторы $\{e_\nu\}$ так, чтобы оператор S имел диагональный вид в соответствующей координатной системе. Так как S имеет конечный след, его собственные значения σ_ν^2 удовлетворяют условию $\sum\limits_{\nu} \sigma_\nu^2 < \infty$. Выберем ξ, как описано, и образуем сумму $y = \sum\limits_{\nu} \xi_\nu e_\nu$, сходящуюся в среднем с вероятностью единица. Используя теорему 6.2.1 (утверждения 3 и 7), получим, что преобразование Фурье распределения суммы y имеет вид

$$\prod_{\nu} \exp\left[-\frac{\sigma_\nu^2}{2}(x_\nu^*)^2\right] = \exp\left[-\frac{1}{2}\sum_{\nu} \sigma_\nu^2 (x_\nu^*)^2\right] =$$

$$= \exp\left[-\frac{1}{2}(Sx^*, x^*)\right],$$

аналогично для $x = m + y$.

Доказательство теоремы 6.3.1. Используем представление элемента x. Тогда получим: $\mathbf{E}x = m$,

$$\mathbf{E}(x^*, x - m)^2 = \sum_{\nu} \mathbf{E}(x_\nu^*, \xi_\nu)^2 = \sum_{\nu} \sigma_\nu^2 (x_\nu^*)^2 = (Sx^*, x^*),$$

что и утверждалось; в обоих соотношениях операции образования математического ожидания и суммирования меняются местами, но законность этого очевидна.

Т е о р е м а 6.3.3. *Если x и y — два нормальных независимых случайных элемента в гильбертовом пространстве со средними значениями соответственно m_x и m_y и ковариационными операторами S_x и S_y, то $x + y$ также имеет нормальное распределение со средним значением $m_x + m_y$ и ковариационным оператором $S_x + S_y$.*

Доказательство очевидно.

Т е о р е м а 6.3.4 (àla Крамèр). *Если x и y — независимые случайные элементы из X и $z = x + y$ нормально, то x и y должны также иметь нормальное распределение.*

Д о к а з а т е л ь с т в о. Нетрудно убедиться, что x и y имеют средние значения и ковариационные операторы; будем рассматривать их как заданные. Для произвольного $x^* \in X$ скалярные случайные величины $x^*(x)$ и $x^*(y)$, как известно, нормальны, и так как их моменты первого и второго порядков могут быть вычислены по фиксированным средним значениям и ковариационным операторам, то их полные распределения определены. Но зная эти распределения, мы можем (как мы заметили раньше) построить распределения вероятностей x и y, и это завершает доказательство (см. замечания $6.3._1$).

Т е о р е м а 6.3.5. *Если B — ограниченное линейное преобразование X в X и x — случайный элемент с нормальным распределением со средним значением m и ковариационным оператором S, то $y = Bx$ имеет нормальное распределение с параметрами*

$$m_y = Bm,$$
$$S_y = BSB^*.$$

Д о к а з а т е л ь с т в о. Для любого x^* имеем $(x^*, y) = (x^*, Bx) = (B^*x^*, x)$. Элемент $y = Bx$ имеет вполне определенное распределение вероятностей. Его преобразование Фурье задается соотношением

$$\mathbf{E} \exp [i (x^*, y)] = \exp [i (B^*x^*, x)] = \exp \left\{ i (B^*x^*, m) - \right.$$
$$\left. - \frac{1}{2} (SB^*x^*, B^*x^*) \right\} = \exp \left\{ i (x^*, Bm) - \frac{1}{2} (BSB^*x^*, x^*) \right\},$$

что доказывает предположение. Можно также непосредственно показать, что $T = BSB^*$ есть S-оператор. Действительно, непосредственно видно, что T — эрмитов, ограниченный, неотрицательно определенный оператор. Он вполне непрерывен, так как он преобразует слабо сходящуюся последовательность в сильно сходящуюся. Для того чтобы показать, что T имеет конечный след, проще всего использовать координатный язык, выбирая систему координат, в которой оператор S имеет диагональную форму, и вычислить след.

Теорема 6.3.6. *Пусть P_n, $n=1, 2, \ldots$, — нормальные распределения в гильбертовом пространстве X со средними значениями m_n и ковариационными операторами S_n. Если $m_n \to m$ (сильно), а $(S_n x^*, x^*) \to (Sx^*, x^*)$ для всех $x^* \in X$ и $S_n \leqslant T$, где T — некоторый S-оператор, то P_n слабо сходится к нормальному распределению со средним m и ковариационным оператором S.*

Доказательство. Мы можем считать, что $m_n = m = 0$, чего можно добиться переносом (или, точнее, сильно сходящейся последовательностью переносов) в гильбертовом пространстве. Для любого $x^* \in X^*$

$$\hat{P}_n(x^*) = \exp\left\{-\frac{1}{2}(S_n x^*, x^*)\right\} \to \exp\left\{-\frac{1}{2}(Sx^*, x^*)\right\} = \hat{P}(x^*).$$

Используем теперь теорему 6.2.3. Имеем

$$E_n x_\nu^2 = \int_X x_\nu^2 P_n(dx) = s_\nu^n,$$

где s_ν^n — $(\nu\nu)$-й элемент оператора S_n, так что

$$E_n \sum_{N+1}^{\infty} x_\nu^2 = \sum_{N+1}^{\infty} s_\nu^n \leqslant \sum_{N+1}^{\infty} t_\nu,$$

где t_ν — $(\nu\nu)$-й элемент оператора T. Следовательно,

$$\lim_{N\to\infty} \sup_n E_n \sum_{N+1}^{\infty} X_\nu^2 \leqslant \lim_{N\to\infty} \sum_{N+1}^{\infty} t_\nu = 0,$$

и условие 1 упомянутой теоремы выполнено, так что результат доказан (см. замечания 6.3.$_2$).

В этом разделе мы изучали нормальные распределения только в гильбертовом пространстве. В общем банаховом пространстве можно было бы по аналогии с гильбертовым пространством определить нормальный случайный элемент x как такой, для которого $x^*(x)$ является нормальной случайной величиной для каждого $x^* \in X^*$. Кое-что из того, о чем было сказано выше, обычно применяется, но пока наши знания о нормальных распределениях в общих банаховых пространствах неполны.

6.4. Закон больших чисел

В начале этого раздела рассмотрим некоторые элементарные ситуации. Пусть $x_1, x_2, x_3, \ldots$ — независимые случайные элементы в банаховом пространстве X с одинаковым распределением вероятности и средним значением m. Образуем среднее

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + \ldots + x_n).$$

Для любого $x^* \in X^*$ имеем

$$x^*(\bar{x}) = \frac{1}{n}[x^*(x_1) + x^*(x_2) + \ldots + x^*(x_n)],$$

и скалярные случайные величины $x^*(x_\nu)$ также независимы и имеют одинаковое распределение. Их среднее значение существует и равно $x^*(m)$. Из усиленного закона больших чисел тогда следует, что случайные величины $x^*(\bar{x})$ сходятся к их среднему значению $x^*(m)$ с вероятностью единица. Используя предположенную сепарабельность пространства X^*, можно выбрать счетную последовательность $x_1^*, x_2^*, \ldots,$ плотную в X^*. Стандартное рассуждение показывает тогда, что почти наверное $\bar{x}$ слабо сходится к m.

Это, однако, нельзя считать существенным обобщением классического закона больших чисел, так как здесь мы имеем дело с асимптотическими свойствами $\bar{x}$, выраженными через x^*. Чтобы перейти к более сильным нормам сходимости, обратимся к одному простому результату, основанному на следующей лемме.

Лемма 6.4.1. (Неравенство Чебышева.) *Пусть x — случайный элемент в гильбертовом пространстве X, для которого* $\mathbf{E}\,\|x\|^2 < \infty$ *и* $\mathbf{E}x = 0$ (второе условие означает просто, что $\mathbf{E}x$ выбирается в начале координат, что достигается переносом в X). *Тогда*

$$P_{\iota}\{\|x\| \geqslant C\} \leqslant \frac{\mathbf{E}\,\|x\|^2}{C^2}\,.$$

Доказательство проводится, как в скалярном случае.

Теорема 6.4.1. (Слабый закон с сильной сходимостью.) *Пусть $x_1, x_2, \ldots, x_n, \ldots$ — некоррелированные случайные элементы в гильбертовом пространстве с общим* $\mathbf{E}\,\|x\|^2 < \infty$. *Обозначим* $\mathbf{E}x = m$. *Тогда для любого* $\varepsilon > 0$

$$P\{\|\bar{x} - m\| > \varepsilon\} \to 0 \quad \textit{при} \quad n \to \infty.$$

Доказательство. Некоррелированность x_ν и x_μ понимается в том смысле, что $\mathbf{E}\,(x_\nu, x_\mu) = (\mathbf{E}x_\nu, \mathbf{E}x_\mu)$. Доказательство ведется очевидным образом. Имеем $\mathbf{E}\,\|\bar{x} - m\|^2 = \frac{1}{n}\,\mathbf{E}\,\|x - m\|^2$, так что лемма дает нам

$$P\{\|\bar{x} - m\| \geqslant \varepsilon\} \leqslant \frac{\mathbf{E}\,\|x - m\|^2}{n\varepsilon^2} \to 0,$$

что и утверждалось.

Если $x_1, x_2, \ldots$ независимы и одинаково распределены, то мы можем использовать преобразование Фурье распределения $\bar{x}$, для которого

$$\hat{P}(x^*) = \left[1 + \frac{ix^*(m)}{n} + O\left(\frac{1}{n^2}\right)\right]^n \to \exp\,[i\,(x^*, m)].$$

Из этой сходимости в силу теоремы 6.2.3 вытекает результат, так как в этом случае

$$R_N \leqslant \sum_N^{\infty} m_\nu^2 + \sum_N^{\infty} \mathbf{E}\,(\xi_\nu - m_\nu)^2 \downarrow 0 \quad \text{при} \quad N \to \infty\,.$$

Здесь m и x записаны в координатной форме:

$$m = (m_1, m_2, \ldots) \quad \text{и} \quad x = (\xi_1, \xi_2, \ldots).$$

Более существен следующий результат.

Теорема 6.4.2. (Усиленный закон с сильной сходимостью.) *Пусть $x_1, x_2, x_3, \ldots$ — независимые одинаково распределенные случайные элементы банахова пространства X и* $\mathbf{E}\|x\| < \infty$. *Тогда $\bar{x}$ сильно сходится с вероятностью единица к среднему значению* $\mathbf{E}x$.

Доказательство. Предположим для начала, что случайные элементы могут принимать только счетное множество значений $\eta_1, \eta_2, \eta_3, \ldots$. Для натурального k определим

$$x_\nu^k = \begin{cases} x_\nu, & \text{если } x_\nu = \eta_1, \eta_2 \ldots \text{ или } \eta_k, \\ 0 & \text{в остальных случаях} \end{cases}$$

и $x_\nu = x_\nu^k + r_\nu^k$. Имеем

$$\bar{x} = \frac{1}{n} \sum_{\nu=1}^{n} x_\nu^k + R_n^k, \quad R_n^k = \frac{1}{n} \sum_{\nu=1}^{n} r_\nu^k.$$

Первая сумма при любом k сходится (сильно) почти наверное к $\mathbf{E}x^k$. Это просто конечномерный вариант классического усиленного закона больших чисел. Также с вероятностью единица имеет место сходимость

$$\frac{1}{n} \sum_{\nu=1}^{n} \| r_\nu^k \| \to \mathbf{E} \| r_\nu^k \|,$$

где правая часть не зависит от ν. Но $\mathbf{E}\| r_\nu^k \| \to 0$ при $k \to \infty$. Комбинируя эти утверждения, мы видим, что $\bar{x}$ сходится почти наверное к $\mathbf{E}x$.

Перейдем теперь к общему случаю. Используя сепарабельность X, выбираем счетную последовательность точек $x_1, x_2, \ldots$, всюду плотную в X. Каждое x_i мы окружаем шарами $S_i(\varepsilon) = \{x \mid \| x - x_i \| \leqslant \varepsilon\}$, так что X покрывается системой перекрывающихся шаров. Введем непересекающиеся множества

$$\begin{aligned} &E_1(\varepsilon) = S_1(\varepsilon), \\ &E_2(\varepsilon) = S_2(\varepsilon) \cap S_1^*(\varepsilon), \\ &E_3(\varepsilon) = S_3(\varepsilon) \cap S_1^*(\varepsilon) \cap S_2^*(\varepsilon), \\ &\ldots\ldots\ldots\ldots \end{aligned}$$

Определим преобразование $x \to T_\varepsilon x$, $T_\varepsilon x = x_i$, где x_i — центр шара $S_i(\varepsilon)$, если $x \in E_i(\varepsilon)$. Ясно, что случайные элементы $T_\varepsilon x_\nu$ обладают теми же свойствами, что и x_ν в первой части доказательства. Следовательно, для любого $\varepsilon > 0$

$$\frac{1}{n}\sum_{\nu=1}^{n} x_\nu = \frac{1}{n}\sum_{\nu=1}^{n} T_\varepsilon x_\nu + \frac{1}{n}\sum_{\nu=1}^{n}(T_\varepsilon x_\nu - x_\nu).$$

Первое слагаемое в правой части сходится почти наверное к $\mathbf{E}\, T_\varepsilon x$, а норма второго слагаемого не превосходит ε. Выбирая теперь последовательность $\varepsilon_1, \varepsilon_2, \ldots \downarrow 0$, легко завершаем доказательство.

Могут представлять интерес и другие типы сходимости, как, например, следующий.

Теорема 6.4.3. (Сходимость в среднем степени α.) *Пусть $x_1, x_2, \ldots$ — последовательность независимых и одинаково распределенных случайных элементов в банаховом пространстве X. Если $\mathbf{E}\,\| x \|^a < \infty$, $1 \leqslant \alpha < \infty$, то*

$$\lim_{n\to\infty} \mathbf{E}\,\| \bar{x} - m \|^\alpha = 0.$$

6.5. Центральная предельная теорема

Линейная структура рассматриваемого пространства X делает возможными формулировку и доказательство аналогов центральной предельной теоремы.

Теорема 6.5.1. *Рассмотрим последовательность $x_1, x_2, x_3, \ldots$ независимых и одинаково распределенных случайных элементов в гильбертовом пространстве X. Предположим, что $\mathbf{E}\,\| x \|^2 < \infty$, и введем среднее значение $\mathbf{E}x = m$ и ковариационный оператор S. Тогда нормированный случайный элемент*

$$y_n = \frac{1}{\sqrt{n}}\sum_{\nu=1}^{n}(x_\nu - m)$$

имеет распределение P_n, сходящееся (слабо) к нормальному распределению в X со средним значением, равным нулю, и ковариационным оператором S.

Доказательство. Используем преобразование Фурье $\hat{P}_n(x^*)$. Для заданного x^* это практически то же самое, что обычная характеристическая функция случайной величины:

$$(x^*, y_n) = \frac{1}{\sqrt{n}} \sum_{\nu=1}^{n} (x^*, x_\nu - m).$$

Но из классического случая известно, что предел характеристической функции этих случайных величин имеет вид

$$\hat{P}_n(x^*) \to \exp\left[-\frac{1}{2}\mathbf{D}(x^*, x_\nu - m)\right] = \exp\left[-\frac{1}{2}(Sx^*, x^*)\right].$$

Чтобы применить теорему 6.2.3, нужно проверить выполнение условия 1) этой теоремы. Записывая $y_n = (y_n^1, y_n^2, \ldots)$, имеем

$$\mathbf{E} \sum_{\nu=N}^{\infty} (y_n^\nu)^2 = \sum_{\nu=N}^{\infty} S_{\nu\nu}$$

(ковариационный оператор должен быть S-оператором), так что

$$\lim_{N\to\infty} R_N = \lim_{N\to\infty} \sum_{\nu=N}^{\infty} S_{\nu\nu} = 0,$$

что доказывает наше утверждение.

Можно также доказать следующую теорему (см. замечания 6.5.).

Теорема 6.5.2. *Пусть $x_1, x_2, x_3, \ldots$ — независимые и одинаково распределенные случайные элементы в рефлексивном G-пространстве с базисом. Предположим, что $\mathbf{E}\,\|x\|^2 < \infty$, и обозначим $\mathbf{E}x = m$. Тогда распределение* *вероятностей нормированной величины y_n (определенной в предыдущей теореме) сходится (слабо) к нормальному распределению.*

6.6. Стохастические распределения Шварца

Пусть Φ — множество всех бесконечно дифференцируемых функций на R^1, равных нулю вне компактных множеств, а D — множество *распределений Шварца* [1]) на R^1. Предполагается, что основные факты относительно этих распределений читателю известны.

Для того чтобы ввести вероятностные меры на D и работать с ними, мы не можем просто применить теорию, изложенную в этой главе. Верно, что D имеет линейную структуру, но оно не является банаховым пространством. Необходимые изменения имеют некоторый интерес и будут коротко здесь описаны.

Пусть $\mathcal{D}$ есть σ-алгебра, порожденная множествами вида $\{d \mid d(\varphi) \leqslant c\} \subset D$, где $\varphi \in \Phi$ и $c \in R^1$ произвольны.

Определение 6.6.1. *Под стохастическим распределением Шварца понимается величина, принимающая значения из D в соответствии с вероятностной мерой, определенной на $\mathcal{D}$.*

Непосредственным следствием этого определения является то, что если d_1 и d_2 — два стохастических распределения Шварца, то произвольная линейная комбинация $c_1d_1 + c_2d_2$ также является стохастическим распределением Шварца. Также если d_n, $n = 1, 2, \ldots$ — последовательность стохастических распределений Шварца, сходящаяся почти наверное (в топологии Шварца), то ее предел также является стохастическим распределением Шварца. Любая производная $D^n d$ стохастического распределения Шварца является стохастическим распределением Шварца, что ясно из соотношения $D^n d(\varphi) = (-1)^n d(D^n \varphi)$.

---

[1]) То есть *обобщенных функций* по принятой в русском языке терминологии.— *Прим. ред.*

Определение 6.6.2. *Говорят, что стохастическое распределение Шварца d имеет среднее значение $m \in D$, если $\mathbf{E}d(\varphi) = m(\varphi)$ для каждого $\varphi \in \Phi$.*

Очевидно, что если d_1 и d_2 имеют средние значения m_1 и m_2, то линейная комбинация $c_1d_1 + c_2d_2$ имеет среднее значение $c_1m_1 + c_2m_2$. Кроме того, если $\mathbf{E}d = m$, то n-я производная d_n распределения d имеет среднее значение $D^n m$; действительно,

$$(-1)^n D^n m(\varphi) = m(D^n\varphi) = \mathbf{E}d(D^n\varphi) = (-1)^n \mathbf{E}D^n d(\varphi),$$

так что $\mathbf{E}D^n d = D^n\mathbf{E}d = D^n m$.

Закон больших чисел может быть сформулирован следующим образом. Рассмотрим последовательность независимых (независимость определяется очевидным образом) и одинаково распределенных стохастических распределений Шварца $d_1, d_2, d_3, \ldots$ со средними значениями $\mathbf{E}d_n = m$. Мы требуем, чтобы для каждой реализации эта последовательность была равностепенно непрерывна в начале координат: для любой последовательности $\varphi_i \in \Phi$, такой, что $\varphi_i \to 0$ при $i \to \infty$, и для любого положительного ε имеем для достаточно больших значений i

$$|d_n(\varphi_i)| \leqslant \varepsilon \text{ для всех } n.$$

Тогда можно получить следующую теорему.

Теорема 6.6.1. *Если d_n независимы, одинаково распределены со средним значением m и каждая реализация этой последовательности равностепенно непрерывна в начале координат, то*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} d_n = m$$

с вероятностью единица.

Относительно доказательства см. замечания $6.6._1$.

Эта теория находится еще в начальном состоянии.

6.7. Примеры

Легко проиллюстрировать понятие математического ожидания в банаховом пространстве. Пусть, например, $X = c$ —

множество всех сходящихся последовательностей $x = (x_1, x_2, \ldots)$ действительных чисел с нормой, соответствующей равномерной сходимости. Предположим, что

$$\mathbf{E}\|x\| = \mathbf{E}\sup_n |x_n| < \infty.$$

Тогда среднее значение случайного элемента x есть просто постоянный элемент m из X с координатами $m = (m_1, m_2, \ldots)$, $m_n = \mathbf{E}x_n$. Так как предел $\lim\limits_n x_n$ существует почти наверное и так как случайная величина $\sup\limits_n x_n$ имеет конечное математическое ожидание, то $\lim\limits_n m_n$ существует, так что $m \in c$. Этот элемент из c должен быть математическим ожиданием x, так как двойственное пространство $c^* = l_1$ состоит из суммируемых последовательностей $y = (y_0, y_1, y_2, \ldots)$, $\sum\limits_0^\infty |y_n| < \infty$. Следовательно, для любого $y \in c^*$ имеем

$$y(m) = y_0 \lim m_n + \sum_1^\infty y_n m_n = \mathbf{E}\left(y_0 \lim x_n + \sum_1^\infty y_n x_n\right) = \mathbf{E}y(x),$$

что и устанавливает справедливость утверждения.

Возьмем теперь $X = l_2$. Если

$$\mathbf{E}\|x\| = \mathbf{E}\sqrt{\sum_1^\infty x_n^2} < \infty,$$

то вектор $m = (m_1, m_2, \ldots)$ с $m_n = \mathbf{E}x_n$ принадлежит l_2; действительно, для любого N имеем

$$\sum_1^N m_n^2 = \mathbf{E}\sum_1^N m_n x_n \leqslant \left(\sum_1^N m_n^2\right)^{1/2} \mathbf{E}\|x\|,$$

так что $\sum\limits_1^\infty m_n^2 < \infty$. Но для любого $y = (y_1, y_2, \ldots) \in l_2^* = l_2$

$$y(m) = \sum_1^\infty y_n m_n = \sum_1^\infty y_n \mathbf{E}x_n = \mathbf{E}\sum_1^\infty y_n x_n = \mathbf{E}y(x),$$

где использовано неравенство Шварца. Следовательно, $\mathbf{E}x = m$.

Наконец, возьмем $X = C\ (0,\ 1)$ — множество непрерывных действительных функций на интервале (0,1) с нормой $\| x\| = \max\limits_{0 \leqslant t \leqslant 1} | x(t)|$, $x = x\ (t)$. Если

$$\mathbf{E}\, \| x \| = \mathbf{E} \max_{0 \leqslant t \leqslant 1} | x(t) | < \infty,$$

то функция $m = m\ (t) = \mathbf{E}x\ (t)$ принадлежит $X = C\ (0,\ 1)$, так как

$$\lim_{h \to 0} | m(t+h) - m(t) | \leqslant \lim_{h \to 0} \mathbf{E}\, | x(t+h) - x(t) | = 0.$$

Так как C^* состоит из ограниченных мер μ на (0, 1), то для любого $\mu \in C^*$ имеем

$$\mu\,(m) = \int_0^1 m\,(t)\,\mu\,(dt) = \int_0^1 \mathbf{E}x\,(t)\,\mu\,(dt) = \mathbf{E} \int_0^1 x(t)\,\mu(dt) = \mathbf{E}\mu\,(x),$$

так что $\mathbf{E}x = m$.

В этой связи можно отметить, что, так как любое сепарабельное банахово пространство изометрично и изоморфно некоторому подпространству пространства $C\ (0,\ 1)$ (это хорошо известная теорема Банаха и Мазура, см. Банах [1]), теория вероятностей на сепарабельном банаховом пространстве имеет дело со строго непрерывными случайными процессами и потому представляет специальный интерес.

Наиболее непосредственными, пожалуй, наиболее важным из приложений этой главы является теория *случайных процессов*. Случайный процесс обычно определяется как функция $x\ (t,\ \omega)$ двух переменных — времени t и случайного параметра ω, при некоторых предположениях измеримости. В линейной теории случайных процессов $x(t,\ \omega)$ обычно рассматривается как множество измеримых функций $x_t\ (\omega)$ (случайных величин) переменной ω, причем каждому значению t соответствует одна такая функция. Если эти функции переменного ω могут быть идентифицированы с элементами некоторого банахова пространства, то процесс представляет собой кривую с параметром t в этом банаховом пространстве. Конечно, можно идти и обратным путем. Для каждого ω мы получаем функцию $x_\omega(t)$ действительного переменного t. Пусть теперь эти функции принадлежат некоторому банахову пространству. Тогда процесс состоит из распределения вероятностей, определенного на банаховом пространстве, что и было исходной точкой настоящей главы.

Пусть $x_\omega(t)$ — случайный процесс, определенный с помощью распределения вероятностей на гильбертовом пространстве $L_2(0, 1)$, причем

$$x_\omega(t) = m(t) + y_\omega(t):$$

Здесь $m(t)$ — фиксированный элемент из $L_2(0, 1)$, $y_\omega = y_\omega(t)$ имеет $\mathbf{E}\,\|y_\omega\|^2 < \infty$ и $\mathbf{E}y_\omega = 0$. Рассмотрим теперь последовательные независимые наблюдения процесса $x_\omega(t)$ и обозначим выборочные функции через $x^1(t)$, $x^2(t)$, ..., $x^n(t)$. Образуем среднее

$$m^* = m^*(t) = \frac{1}{n} \sum_{\nu=1}^{n} x^\nu(t),$$

также являющееся элементом гильбертова пространства. Тогда из теоремы 6.4.2 известно, что, если исключить событие вероятности нуль, оценка m^* сильно сходится к m при n, стремящемся к бесконечности. Другими словами, функция $m(t)$ имеет состоятельную оценку — среднее $m^*(t)$. Используя теорему 6.5.1, можно изучить сходимость распределений среднего m^*. Очевидно, что этот подход полезен при статистическом изучении стохастических процессов.

Рассмотрим теперь нечто аналогичное. Если $F(x)$ — непрерывная функция распределения действительной случайной величины x, то делается обычное преобразование $y = F(x)$ к равномерному распределению. Имея выборку $y_1, y_2, \ldots, y_n$ из n наблюденных значений, образуем эмпирическую функцию распределения

$$F^*(u) = \frac{\text{Число значений } y_\nu \leqslant u}{n} = \frac{1}{n} \sum_{\nu=1}^{n} \varepsilon_{y_\nu}(u),$$

где $\varepsilon_y(u)$ равно 1, если $y \leqslant u$, и 0 в противоположном случае. Величины $\varepsilon_y(u)$ являются случайными величинами в $L_2(0, 1)$ и

$$\mathbf{E}\varepsilon(u)=u,$$

$$\mathbf{E}\|\varepsilon(u)\|^2=\frac{1}{2}<\infty.$$

Таким образом, $F^*(u)\in L_2(0, 1)$ сходится почти наверное по норме к истинной функции распределения u. Из теоремы 6.5.1 можно вывести, что если $f(z)$ — непрерывный функционал, определенный в $L_2(0, 1)$ и принимающий действительные значения, то распределение случайной величины

$$f_n=f\left(\frac{F^*(u)-u}{\sqrt{n}}\right)$$

сходится к распределению величины $f(Z)$, где Z есть случайный элемент в гильбертовом пространстве, имеющий нормальное распределение. Среднее значение величины Z равно нулю, а ее ковариационный оператор S определяется соотношением

$$(Sz^*, z^*)=\int_0^1\int_0^1 z^*(u)\,z^*(v)\,K(u, v)\,du\,dv,$$

где $z^*=z^*(u)\in L_2(0, 1)$ и $K(u, v)=\min(u, v)-uv$.

Сходимость эмпирических функций распределения может быть изучена более общим образом. Мы можем использовать другие топологии при установлении сходимости. Можно попытаться также перенести этот результат на более общие пространства, чем действительная прямая. В этом направлении были получены некоторые результаты (см. замечания 6.7). Пусть S — полное и сепарабельное пространство с метрикой $d(s', s'')$. Вероятностная мера P определена на борелевских множествах из S, и мы требуем, чтобы $\int_S d(s_0, s)\,P(ds)<\infty$. Произведем n независимых наблюдений $s_1, s_2, \ldots, s_n$ с распределением P и образуем вероятностную меру

$$P^*=\frac{1}{n}\sum_{\nu=1}^{n}\delta_{s_\nu}.$$

Как и выше, можно сказать, что P^* аппроксимирует P, если n достаточно велико, и это может быть следующим образом превращено в утверждение относительно состоятельной оценки. Пусть F — множество действительных функций $f(s)$ на S, таких, что

$$\sup \frac{|f(s')-f(s'')|}{d(s', s'')} < \infty .$$

Очевидным образом вводя в F норму, получим банахово пространство. Можно убедиться в том, что множество функций вида $P^* - P$ принадлежит некоторому подпространству Φ банахова пространства F^*, и мы можем теперь применить различные формулировки закона больших чисел из разд. 6.4. Отсюда следует, например, что почти наверное $\| P^* - P\| \to 0$ при $n \to \infty$ по норме в Φ. В частном случае $S = R^1$ получаем

$$\lim_{h\to\infty} \int_{-\infty}^{\infty} |F^*(y) - F(y)|\, dy = 0$$

при условии

$$\int_{-\infty}^{\infty} |y|\, dF(y) < \infty .$$

Глава 7

СТОХАСТИЧЕСКИЕ АЛГЕБРЫ

7.1. Аддитивные и мультипликативные предельные теоремы

В последней главе мы предполагали, что наши случайные элементы принимают значения из сепарабельного банахова пространства X, и использовали линейные свойства пространства X для формулировки и доказательства результатов. Теперь мы предположим, что X образует сепарабельную *топологическую алгебру*. В этой алгебре установлена непрерывная бинарная (не обязательно коммутативная) операция умножения xy, x и $y \in X$. Пусть x и y — независимые случайные элементы из X; тогда можно образовать новый случайный элемент $z = xy$. Обозначая соответственные распределения вероятностей P_x, P_y, и P_z, мы видим (см. замечания 7.1.$_1$), что P_z есть вполне определенная вероятностная мера. Будем писать $P_z = P_x \circ P_y$. Ясно, что мы можем изучать задачи, выраженные в терминах $\circ$-операции, так же как мы изучали аналогичные задачи в терминах $*$-операции.

Более существенно, пожалуй, выяснить логические взаимоотношения между этими двумя типами результатов. Для определенности укажем на одну особенно важную задачу. Если $P_1^{(n)}$, $P_2^{(n)}$, . . ., $P_n^{(n)}$ — вероятностные меры, принадлежащие $\mathcal{P}(X)$, то можно образовать два новых распределения вероятностей

$$Q_n = P_1^{(n)} * P_2^{(n)} * \cdots * P_n^{(n)},$$

$$R_n = P_1^{(n)} \circ P_2^{(n)} \circ \cdots \circ P_n^{(n)}.$$

Если можно установить предельные законы для одного из этих распределений Q_n или R_n, то что можно утверждать относительно другого? Часто проще иметь дело с Q_n, так как сложение в алгебре коммутативно, и это возвращает нас к задачам в банаховом пространстве, подобным изученным в гл. 6. В этом разделе мы докажем несколько полезных соотношений указанного типа.

Начнем с терминологии. Рассмотрим случайный процесс $x(t)$, $0 \leqslant t < \infty$, принимающий значения из X и имеющий распределения вероятностей $P_t \in \mathcal{P}(X)$. Этот процесс должен иметь независимые «приращения» (в аддитивном или мультипликативном смысле), и распределение вероятностей приращения должно зависеть только от длины соответствующего интервала времени. Если имеет место случай независимых приращений, такой, что $P_s * P_t = P_{s+t}$; $s, t \geqslant 0$, то $\{P_t\}$ или $x(t)$ называется *однородным аддитивным процессом*. Если мы имеем дело с независимыми мультипликативными приращениями, такими, что $P_s \circ P_t = P_{s+t}$, $s, t \geqslant 0$, то $\{P_t\}$ или $x(t)$ называется *однородным мультипликативным процессом*. В дополнение мы наложим еще на процесс некоторое условие непрерывности.

Предположим сначала, что пространство X — банахова алгебра, так что его норма удовлетворяет условие. $\| xy \| \leqslant \| x \| \cdot \| y \|$, и X имеет единичный элемент e. Пусть $y(t)$ — однородный аддитивный процесс, принимающий значения из X и удовлетворяющий условию непрерывности

$$C_1 : \sum_{\nu} \mathbf{E} \| y(t_\nu) - y(t_{\nu-1}) \| \leqslant M < \infty$$

для любого разбиения $t_0 = 0,\ t_1,\ t_2,\ \ldots,\ t_{n-1},\ t = t_n$ заданного интервала $(0, t)$. Построим теперь однородный мультипликативный процесс $x(t)$, связанный с $y(t)$ естественным образом. Выпишем формальное выражение (см. замечания 7.1.$_2$)

$$x(t) = e + \int_0^t dy(s) + \iint\limits_{0<s_1<s_2<t} dy(s_1)\, dy(s_2) + \\ + \iiint\limits_{0<s_1<s_2<s_3<t} dy(s_1)\, dy(s_2)\, dy(s_3) + \ldots,$$

которому мы придадим сейчас точное значение.

Эти интегралы будут пониматься как пределы в сильной $L_1(X)$-топологии обычных сумм Римана — Стильтьеса. Рассмотрим второй (двойной) интеграл, который типичен.

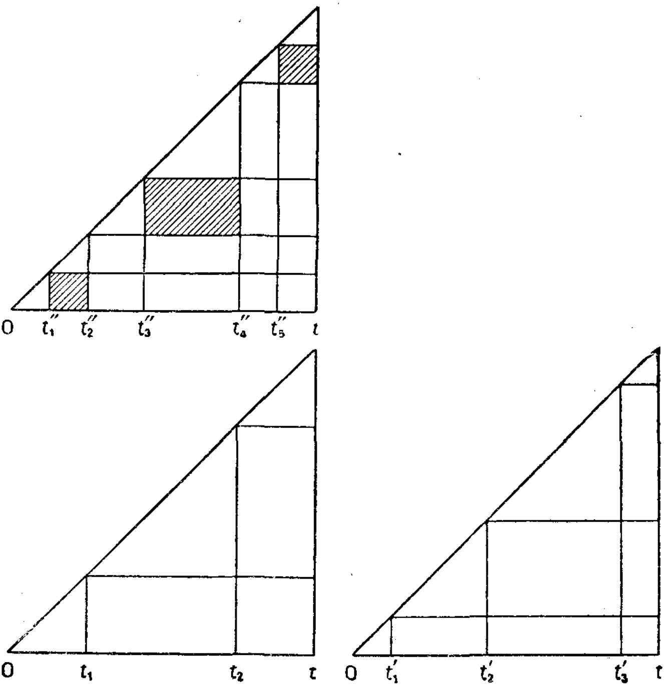

Рис. 1

Для разбиения точками $\{t_\nu\}$ выпишем сумму Римана — Стильтьеса

$$S = \sum_{\nu<\mu} [y(t_\nu) - y(t_{\nu-1})]\,[y(t_\mu) - y(t_{\mu-1})],$$

где порядок множителей может быть существен. При измельчении разбиения эти суммы будут сходиться к пределу, который не зависит от того, какая последовательность разбиений выбрана. Для того чтобы доказать это, возьмем другое разбиение $t_0' = 0 < t_1' < t_2' < \ldots < t_{m-1}' < t = t_m'$ и обозначим соответствующую сумму через S'. Объединенное разбиение $t_0'' = 0 < t_1'' < t_2'' < \ldots < t$ приводит к сумме S'' (рис. 1).

Каждому прямоугольнику на рисунке соответствует одно слагаемое в соответствующей сумме. В разности $S'' - S$ участвуют только заштрихованные прямоугольники, что следует из тождества $(a+b)(c+d)=ac+bc+bd+ad$. Теперь ясно, что́ имеет место в общем случае: разность $S''-S$ состоит из прямоугольников, примыкающих к диагонали, и ее норма удовлетворяет неравенству

$$\|S''-S\|\leqslant\sum_{\nu}\|[y(t''_{\nu})-y(t''_{\nu-1})]\,[y(t''_{\nu+1})-y(t''_{\nu})]\|\leqslant$$
$$\leqslant\sum_{\nu}\|z(t''_{\nu})-y(t''_{\nu-1})\|\cdot\|y(t''_{\nu+1})-y(t''_{\nu})\|,$$

так что $L_1(X)$-норма разности удовлетворяет неравенству

$$\mathbf{E}\|S''-S\|\leqslant\sum_{\nu}\mathbf{E}\|y(t''_{\nu})-y(t''_{\nu-1})\|\cdot\mathbf{E}\|y(t''_{\nu+1})-y(t''_{\nu})\|$$

и стремится к нулю при $\max(t''_{\nu+1}-t''_{\nu})\to 0$. Таким образом, S'' произвольно мало отличается от S, а следовательно, и от S', что показывает, что двойной интеграл однозначно определен. Но мы получаем также границу для $L_1(X)$-нормы интегралов, рассматривая суммы Римана

$$\mathbf{E}\|\int\int\ldots\int_{0<s_1<s_2<\ldots<s_k<t} dy(s_1)\,dy(s_2)\ldots dy(s_k)\|\leqslant\frac{M^k}{k!}.$$

Следовательно, сумма, определяющая $x(t)$, сходится в сильной $L_1(X)$-топологии, и мы имеем однородный мультипликативный процесс, так как

$$x(t+h)=x(t)\left\{e+\int_t^{t+h}dy(s)+\int\int_{t<s_1<s_2<t+h}dy(s_1)\,dy(s_2)+\ldots\right\}.$$

Для того чтобы это проверить, выписывается соответствующее соотношение для аппроксимирующих сумм и совершается переход к пределу. Ясно, что выражение в скобках не зависит от $x(t)$ и имеет то же распределение, что и $x(h)$. Мы могли бы записать символически

$$dx(t)=x(t)\,dy(t)$$

нли

$$x(t)=e+\int_0^t x(s)\,dy(s).$$

Ясно также, что процесс x непрерывен.

Теорема 7.1.1. *Пусть $y(t)$ — однородный случайный процесс, принимающий значения из банаховой алгебры X с единичным элементом e и удовлетворяющий условию непрерывности C_1. Тогда*

$$x(t)=e+\int_0^t dy(s)+\iint\limits_{0<s_1<s_2<t} dy(s_1)\,dy(s_2)+\ldots$$

определяет однородный мультипликативный процесс.

Можно заметить, что если X — коммутативная алгебра, то мы имеем просто

$$x(t)=e+\int_0^t dy(s)+\frac{1}{2}\int_0^t\int_0^t dy(s_1)\,dy(s_2)+$$

$$+\frac{1}{6}\int_0^t\int_0^t\int_0^t dy(s_1)\,dy(s_2)\,dy(s_3)+\ldots=\exp\int_0^t dy(s)=\exp y(t).$$

Эта теорема представляется привлекательной, но ее польза ограничена условием непрерывности C_1. Предположим теперь, что X — банахово пространство и топологическая алгебра с единичным элементом и что $y(t)$ — однородный аддитивный процесс со значениями из X, такой, что $\mathbf{E}\|y(t)\|^2<\infty$, так что можно использовать топологию $L_2(X)$. Для того чтобы определить $x(t)$, мы исходим из сумм Римана—Стильтьеса того же вида. В области $S_r\subset R^r$ всех точек $(t_1, t_2, \ldots, t_r)$, таких, что $0<t_1<$ $<t_2<\ldots<t_r<t$, мы получаем параллелепипеды со сторонами $I_1, I_2, \ldots, I_r$, параллельными осям координат, или суммы таких параллелепипедов. С подобным параллелепипедом ϱ мы связываем случайный элемент $y(\varrho)=\Delta y(I_1)\,\Delta y(I_2)\ldots\Delta y(I_r)$, где порядок сомножителей может быть существенным. Заметим, что таким путем мы можем определить аддитивную случайную функцию множества $y(\sigma)$, где σ — конечная сумма параллелепипедов, но эта функция множества не обладает, вообще говоря, независимыми приращениями. Нам нужна граница для $\mathbf{E}\|y(\sigma)\|^2$, и мы будем предполагать, что выполнено условие

$$C_2: \mathbf{E}\|y(\sigma)\|^2\leqslant C^r\cdot\text{объем }\sigma,$$

где C — некоторая константа, а под объемом понимается лебегова мера в r-мерном пространстве. Теперь мы можем поступать так же, как и раньше. Интересующие нас стохастические интегралы однозначно определены. Действительно, разность между различными суммами Римана — Стильтьеса может быть выражена, как раньше, через значения $y(\sigma)$, где элементарные параллелепипеды, образующие σ, покрывают только небольшую часть S_r, и отсюда вытекает единственность, как и в предыдущем случае. $L_2(X)$-норма такого интеграла ограничена:

$$\mathbf{E}\,\|J_r(t)\|^2 = \mathbf{E}\left\|\int_{S_r} dy(t_1)\,dy(t_2)\,\ldots\,dy(t_r)\right\|^2 \leqslant C^r \frac{t^r}{r!},$$

так что

$$\sum_{r=0}^{\infty} \sqrt{\mathbf{E}\,\|J_r(t)\|^2} \leqslant \sum_{r=0}^{\infty} \frac{(Ct)^{r/2}}{\sqrt{r!}} < \infty.$$

Отсюда вытекает сходимость в сильной $L_2(X)$-топологии рядов

$$x(t) = \sum_{r=0}^{\infty} J_r(t), \quad J_0(t) = e.$$

Вернемся на минуту к новому условию непрерывности C_2. Его выполнение приходится проверять, но можно заметить, что если норма в X вводится с помощью скалярного произведения (x, y), то C_2 может быть несколько упрощено. Действительно, если $R' = I_1' \times I_2' \times \ldots \times I_r'$ и $R'' = I_1'' \times I_2'' \times \ldots \times I_r''$ — два непересекающихся параллелепипеда в S_r, то

$$\mathbf{E}(y(R'),\ y(R'')) = \mathbf{E}(y(I_1')\,y(I_2')\,\ldots\,y(I_r'),\ y(I_1'')\,y(I_2'')\,\ldots\,y(I_r'')).$$

По крайней мере для некоторого i должно иметь место $I_i' \cap I_i'' = \phi$. Если теперь ожидаемое значение процесса y равно нулю, то указанное скалярное произведение равно нулю. В таком случае нужно только проверить выполнение условия C_2 для произвольного параллелепипеда σ.

Теорема 7.1.2. *Пусть $y(t)$ — однородный мультипликативный процесс, принимающий значения из банахова* *пространства X, образующего топологическую алгебру с единичным элементом. Если $y(t)$ удовлетворяет условию непрерывности C_2, то однородный мультипликативный процесс $x(t)$ может быть единственным образом определен, как было указано выше.*

Теперь мы можем выводить мультипликативные предельные теоремы из соответствующих аддитивных теорем. Рассмотрим треугольную последовательность случайных элементов, принимающих значения из X,

$$\begin{array}{l} y_{11}, \\ y_{21},\ y_{22}, \\ y_{31},\ y_{32},\ y_{33}, \\ \dots\dots\dots\dots \end{array}$$

В каждой строке элементы предполагаются независимыми и одинаково распределенными. Как и раньше, мы будем иметь дело с двумя случаями, соответствующими методам $L_1(X)$ и $L_2(X)$. Нам опять понадобится условие непрерывности, но выраженное в терминах величин $y_{n\nu}$. В суммах, фигурирующих в условиях C_1 и C_2, заменим множители вида $y(t_\nu) - y(t_{\nu-1})$ величинами $y_{n\nu}$ и выпишем соответствующие условия C_1' и C_2'. Сначала возьмем простую форму

$$C_1': \qquad \sum_{\nu=1}^{n} \mathbf{E}\, \| y_{n\nu} \| \leqslant M < \infty,$$

так как вторая выписывается более сложно.

Теорема 7.1.3. *Пусть $y_{n\nu}$ принимают значения из банаховой алгебры X и удовлетворяют условию C_1'. Пусть $y(t)$ — аддитивный однородный процесс, удовлетворяющий условию C_1, и такой, что для каждого c между 0 и 1 распределения величин*

$$\eta_n = \sum_{\nu=1}^{[cn]} y_{n\nu}$$

слабо сходятся к распределению величины $y(c)$. Тогда распределения величин

$$\xi_n = (e + y_{n1})(e + y_{n2}) \ldots (e + y_{nn})$$

слабо сходятся к распределению величины $x(1)$, где $x(t)$ — однородный мультипликативный процесс, связанный с $y(t)$.

Доказательство. Имеем

$$\begin{cases} \xi_n = e + S_1^{(n)} + S_2^{(n)} + \ldots + S_n^{(n)}, \\ x(1) = e + S_1 + S_2 + \ldots + S_n + \ldots, \end{cases}$$

где

$$\begin{cases} S_1^{(n)} = \sum_{\nu=1}^{n} y_{n\nu}, \\ S_2^{(n)} = \sum_{1 \leqslant \nu < \mu \leqslant n} y_{n\nu} y_{n\mu}, \\ \ldots\ldots\ldots\ldots\ldots\ldots \end{cases}$$

и

$$\begin{cases} S_1 = \int_0^1 dy(t) = y(1), \\ S_2 = \iint_{0<s_1<s_2<1} dy(s_1)\, dy(s_2), \\ \ldots\ldots\ldots\ldots\ldots\ldots \end{cases}$$

Известно, что при использовании L_1-топологии на X произойдут только незначительные изменения, если оборвать ряд для $x(1)$ на индексе суммирования $n = n_0$ при достаточно большом n_0. Мы знаем также, что если заменить каждое слагаемое этой конечной суммы суммой Римана — Стильтьеса, то происходящее от этого изменение невелико, если разбиение интервала (0, 1) взято достаточно мелким. То же самое можно проделать с выражением для ξ_n: оборвать сумму и каждое слагаемое в новой сумме заменить суммой, соответствующей разбиению $0 < [nt_1] < < [nt_2] < \ldots < n$. Но распределение каждой такой суммы слабо сходится к соответствующей величине, образованной приращениями процесса y. Используя независимость величин $y_{n\nu}$ (и приращений процесса y), получаем высказанный результат.

Мы можем использовать это же рассуждение для того, чтобы получить следующий альтернативный результат.

Теорема 7.1.4. *Пусть величины $y_{n\nu}$ принимают значения из банахова пространства и топологической алгебры X и удовлетворяют условию C'_2. Пусть $y(t)$ — аддитивный однородный процесс, удовлетворяющий условию C_2. Тогда выполнено утверждение предыдущей теоремы.*

Две последние теоремы имеют совершенно общий характер. Более специальным результатом является следующий.

Теорема 7.1.5. (Мультипликативный закон больших чисел.) *Пусть y_ν — независимые, одинаково распределенные элементы банаховой алгебры, такие, что $\mathbf{E}\,\|y_1\| < \infty$. Тогда*

$$\gamma_n = \left(e + \frac{1}{n} y_1\right)\left(e + \frac{1}{n} y_2\right) \ldots \left(e + \frac{1}{n} y_n\right)$$

сильно сходится по вероятности к элементу

$$\gamma = \exp m, \text{ где } m = \mathbf{E} y_1.$$

Доказательство. Имеем

$$\gamma_n = e + S_1^{(n)} + S_2^{(n)} + \ldots + S_n^{(n)},$$

где

$$\begin{cases} S_1^{(n)} = \dfrac{1}{n} \sum\limits_{1 \leqslant \nu \leqslant n} y_\nu, \\ S_2^{(n)} = \dfrac{1}{n^2} \sum\limits_{1 \leqslant \nu < \mu \leqslant n} y_\nu y_\mu \\ \text{и т. д.} \end{cases}$$

Но мы знаем, что с вероятностью единица (см. разд. 6.4) имеет место сильная сходимость $S_1^{(n)} \to m$. Рассмотрим

$$S_2^{(n)} = \frac{1}{n} \sum_{\mu=1}^{n} \frac{\mu - 1}{n} S_1^{(\mu)} y_\mu,$$

где

$$S_1^{(\mu)} = \frac{1}{\mu - 1} \sum_{\nu=1}^{\mu-1} y_\nu.$$

Но с вероятностью единица

$$S_1^{(\mu)} = m + \varepsilon_\mu, \quad \|\varepsilon_\mu\| \to 0,$$

так что

$$S_2^{(n)} = \frac{m}{n} \sum_{\mu=1}^{n} \frac{\mu-1}{n} y_\mu + \frac{1}{n} \sum_{\mu=1}^{n} \frac{\mu-1}{n} \varepsilon_\mu y_\mu.$$

Следовательно, с вероятностью единица

$$S_2^{(n)} = \frac{m^2}{2} + \delta_n, \quad \| \delta_n \| \to 0.$$

Вообще с вероятностью единица имеет место сильная сходимость

$$S_n^{(k)} \to \frac{m^k}{k!}\,.$$

Теперь надо дополнить это рассуждение простым использованием равномерности. Имеем

$$\| S_\nu^{(n)} \| \leqslant \frac{1}{n^\nu} \sum_{1 \leqslant k_1 \leqslant k_2 < \ldots < k_\nu \leqslant n} \| y_{k_1} \| \cdot \| y_{k_2} \| \ldots \| y_{k_\nu} \|,$$

так что

$$\mathbf{E} \| S_\nu^{(n)} \| \leqslant \frac{(\mathrm{E} \| y_1 \|)^\nu}{\nu!}\,.$$

Комбинируя это со сказанным выше, мы видим, что для любого $\varepsilon > 0$ имеет место сходимость $\mathbf{P}\{\|\gamma_n - \gamma\| > \varepsilon\} \to 0$ при $n \to \infty$, где γ — элемент

$$e + m + \frac{m^2}{2} + \frac{m^3}{6} + \ldots = \exp m \in X$$

(см. замечания 7.1.$_3$).

7.2. Вероятности на банаховых алгебрах

Вместе с распределением вероятностей P на банаховой алгебре X можно изучать свойства производных величин и понятий в этой алгебре. Стандартные алгебраические и аналитические операции, определенные на X, приводят к случайным элементам различной степени сложности. Распределения вероятностей этих элементов можно было бы определить, вычислить и изучить. Однако при попытке это сделать мы сталкиваемся со многими трудностями, и здесь будет дан только набросок начал этой теории.

Пусть x — случайный элемент в X. Рассмотрим множество R регулярных элементов алгебры, т. е. тех элементов, которые обладают обратным. R есть открытое множество (см. замечания 7.2.$_1$), откуда вытекает, что событие, заключающиеся в том, что x — регулярный элемент, является борелевским множеством, так что $P\ (R)$ определено. Если $P\ (R) > 0$, то можно ввести условное распределение вероятностей для x^{-1} при условии $x \in R$ соотношением

$$P\{x^{-1} \in A \mid x \in R\} = \frac{P\{x^{-1} \in A,\ x \in R\}}{P\ (x \in R)},$$

где A — борелевское множество в X. Но x^{-1} есть непрерывная функция от x при $x \in R$, так что событие $\{x^{-1} \in A, x \in R\}$ является борелевским множеством, и можно говорить о распределении обратной величины x^{-1}.

Для фиксированного значения λ можно определить обычным образом резольвенту $R\ (\lambda,\ \gamma) = (\lambda e - x)^{-1}$, если этот обратный элемент существует. Множество значений λ, для которых резольвента не существует, есть спектр элемента x. В рассматриваемом случае спектр есть случайное множество S. Представляется правдоподобным, что понятие *случайного спектра* будет играть важную роль в будущем развитии этой теории, и следовало бы уделить некоторое внимание его удовлетворительному определению. Конечно, можно для фиксированного λ говорить о вероятности

$$P(\lambda) = P\{\lambda \in S\} = P\{\lambda e - x \bar{\in} R\},$$

и эта функция определяет некоторое среднее поведение спектра. Введем индикатор $I\ (\lambda,\ x)$ множества $\{x \mid \lambda e - x$ сингулярно$\} \subset X$. Имеем $P\ (\lambda) = E\ \{I\ (\lambda,\ x)\}$. Для любой константы c множество $\{(\lambda,\ x) \mid I\ (\lambda,\ x) < c\}$ является открытым множеством в пространстве $\Lambda \times X$, где Λ — комплексная λ-плоскость. Отсюда вытекает, что функция $I\ (\lambda,\ x)$ измерима по Борелю на произведении пространств $\Lambda \times X$, и можно применить теорему Фубини. В частности, интегралы Лебега вида

$$\int\limits_{\lambda \in L} I\ (\lambda,\ x)\, d\lambda$$

по борелевским множествам $L \subset \Lambda$ конечной лебеговой меры существуют почти наверное и являются случайными величинами. Нами доказана

Теорема 7.2.1. *Пусть x — случайный элемент в банаховой алгебре. Множество R регулярных элементов имеет вполне определенную вероятность $P(R)$. Если $P(R) > 0$, то можно говорить об условном распределении вероятностей x^{-1} для x из R. Индикатор $I(\lambda, x)$ случайного спектра S измерим по Борелю на произведении пространств $\Lambda \times X$.*

Для некоторых случайных спектров может быть определена *функция мощности* $s(\lambda)$. Если спектры являются счетными точечными множествами на действительной прямой, такими, что математическое ожидание числа собственных значений в любом заданном интервале (α, β) существует и равно $F(\beta) - F(\alpha)$, где $F(x)$ — дифференцируемая функция, то производная $F'(x)$ называется функцией мощности в силу ее интерпретации в терминах квантовой механики. Мы вернемся к этому понятию в разд. 7.5.3.

Теперь мы можем пойти дальше и изучить другие величины, связанные со случайными спектрами. Так, спектральный радиус

$$r(x) = \sup_{\lambda \in S} |\lambda|,$$

или, что эквивалентно,

$$r(x) = \lim_{n \to \infty} \| x^n \|^{1/n},$$

есть вполне определенная случайная величина. Простейший способ убедиться в этом состоит, пожалуй, в замечании, что $\| x^n \|^{1/n}$ является случайной величиной для любого конечного n, и, как известно, эти случайные величины сходятся при любом $x \in X$.

Это приводит нас к соответствующей проблеме для произведений случайных элементов в банаховой алгебре. Предположим, что $x_1, x_2, \ldots, x_n$ — независимые и одинаково распределенные случайные элементы в X, и образуем произведение

$$y_n = x_1 x_2 \ldots x_n.$$

Что можно сказать о норме произведения $\|\gamma_n\|$ для больших значений n (см. замечания 7.2.$_2$)?

Т е о р е м а 7.2.2. *Предположим, что $x_1, x_2, \ldots, x_n$ — независимые и одинаково распределенные случайные элементы и $\mathbf{E}\log^+\|x_1\| < \infty$. Тогда предел*

$$r = \lim_{n\to\infty} \frac{1}{n}\mathbf{E}\log\|x_1x_2\ldots x_n\|$$

существует и $-\infty \leqslant r < \infty$. Если $r \neq -\infty$, то почти наверное

$$r = \overline{\lim_{n\to\infty}} \frac{1}{n}\log\|x_1x_2\ldots x_n\|.$$

Д о к а з а т е л ь с т в о. Рассмотрим последовательность

$$\alpha_n = \mathbf{E}\log\|x_1x_2\ldots x_n\|.$$

Либо все α_n конечны, либо все они равны $-\infty$, начиная с некоторого n_0. В первом случае можно использовать свойство полуаддитивности

$$\alpha_{n+m} = \mathbf{E}\log\|x_1x_2\ldots x_{n+m}\| \leqslant \mathbf{E}\log\{\|x_1\ldots x_n\|\cdot\|x_{n+1}\ldots \ldots x_{n+m}\|\} = \mathbf{E}\log\|x_1\ldots x_n\| + \mathbf{E}\log\|x_{n+1}\ldots x_{n+m}\| = \alpha_n + \alpha_m$$

и тот факт, что $(1/n)\alpha_n$ имеет предел (см. замечания 7.2.$_3$). Этим доказано первое утверждение. Для завершения доказательства введем величины

$$\xi_n = \frac{1}{n}\log\|x_1x_2\ldots x_n\|.$$

Если n имеет вид $r = \nu\mu$, то

$$\xi_n = \xi_{\nu\mu} = \frac{1}{\nu\mu}\log\|x_1\ldots x_\nu x_{\nu+1}\ldots x_{\nu\mu}\| \leqslant$$

$$\leqslant \frac{1}{\mu}\Big[\frac{1}{\nu}\log\|x_1\ldots x_\nu\| + \frac{1}{\nu}\log\|x_{\nu+1}\ldots x_{2\nu}\| + \ldots \ldots + \frac{1}{\nu}\log\|x_{(\mu-1)(\nu+1)}\ldots x_{\mu\nu}\|\Big] \to$$

$$\to \frac{1}{\nu}\mathbf{E}\log\|x_1\ldots x_\nu\|$$

почти наверное при $\mu \to \infty$. Правая часть близка к r для достаточно большого ν. С помощью стандартного рассуждения можно справиться и со значениями n, не имеющими вида $\nu\mu$, и убедиться, что почти наверное $\overline{\lim\limits_n}\, \xi_n \leqslant r$. Введем теперь функцию

$$f_n = \frac{1}{n} \sum_1^n \log \| x_i \| - \frac{1}{n} \log \| x_1 \ldots x_n \|.$$

Так как $\| x_1 \ldots x_n \| \leqslant \| x_1 \| \ldots \| x_n \|$, то величины f_n неотрицательны. Но

$$\mathbf{E} f_n = \mathbf{E} \log \| x_1 \| - \frac{1}{n} \alpha_n \to \mathbf{E} \log \| x_1 \| - r,$$

и можно применить лемму Фату (см. замечания 7.2.$_4$). Из нее вытекает, что величина $f = \overline{\lim\limits_n}\, f_n = \mathbf{E} \log \| x_1 \| - \overline{\lim\limits_n}\, \xi_n$ имеет конечное математическое ожидание и что

$$\mathbf{E} f = \mathbf{E} \log \| x_1 \| - \mathbf{E}\, \overline{\lim_n}\, \xi_n \leqslant \overline{\lim_n}\, \mathbf{E} f_n = \mathbf{E} \log \| x_1 \| - r,$$

так что $\mathbf{E} \lim\limits_n \xi_n \geqslant r$. Но уже известно, что с вероятностью единица $\overline{\lim\limits_n}\, \xi_n \leqslant r$, так что $\overline{\lim\limits_n}\, \xi_n = r$ почти наверное, что и завершает доказательство.

7.3. Стохастические операторы и случайные уравнения

Пусть Z—банахово пространство. Рассмотрим множество X всех ограниченных линейных преобразований пространства Z в Z. С операторной нормой множество X образует банахову алгебру. Введя в X вероятностную меру P, как в предыдущем разделе, естественно говорить об элементах X как о *стохастических операторах*. Можно изучать обратный элемент x^{-1}, если он существует, спектральные свойства операторов x и т. д.

Более общим образом можно исходить из *полного метрического пространства* Z, в котором введена метрика $\varrho(z_1, z_2)$, и рассматривать множество X преобразований x пространства Z в Z. Если ввести вероятностную меру в X так, чтобы для любого заданного $z \in Z$ элемент xz был измерим по Борелю, то снова можно говорить о стохастическом операторе, хотя мы и можем выйти за пределы банаховой алгебры. При изучении резольвент и т. п. необходимо доказательство различных утверждений измеримости, но это не будет рассматриваться в тексте (ссылки см. в замечаниях).

Важный частный случай последнего типа получаем, когда X состоит из равномерно сжатых преобразований, т. е. таких, что для некоторого c, $0 < c < 1$, и для всех $x \in X$ выполняется неравенство

$$\varrho(xz_1, \ xz_2) \leqslant c\varrho(z_1, \ z_2).$$

Рассмотрим *случайное уравнение* $xz = z$. Имеет место следующий стохастический вариант *банаховой теоремы о неподвижной точке.*

Теорема 7.3.1. *Пусть X состоит из равномерно сжатых преобразований полного сепарабельного метрического пространства Z в Z. На X имеем вероятностную меру, такую, что xz есть измеримый по Борелю случайный элемент. Тогда существует однозначно определенный измеримый по Борелю случайный элемент ξ со значениями из Z, удовлетворяющий случайному уравнению $x\xi = \xi$. Это решение может быть получено итерационным процессом, исходя из произвольного измеримого по Борелю случайного элемента ξ_1:*

$$\begin{gathered} x\xi_1 = \xi_2, \\ x\xi_2 = \xi_3, \\ \cdots\cdots \\ x\xi_n = \xi_{n+1}, \\ \cdots\cdots \end{gathered}$$

Тогда почти наверное $\xi = \lim\limits_{n} \xi_n$.

Доказательство. Существен только вопрос об измеримости. Пусть $z_1, \ z_2, \ \ldots$ есть всюду плотная последовательность в Z. Образуем, исходя из сферических окрестностей, подмножества

$$A_{in} = \left\{ z \mid \varrho(z, z_i) \leqslant \frac{1}{n} \right\} \cap \left\{ \bigcup_{j=1}^{i-1} A_{jn} \right\}^*; \quad i = 1, \ 2, \ \ldots \ .$$

Определим теперь стохастический оператор $x_n z$, где z — заданный случайный элемент, соотношением

$$x_n z = x z_i, \quad \text{если} \quad z \in A_{in}.$$

Но отсюда вытекает, что событие, заключающееся в том, что $x_n z$ принимает значение из борелевского множества B, есть сумма событий $x z_j \in B$ (если $z \in A_{jn}$) при $j = 1, 2, \ldots$. Заметим, что события A_{jn} не перекрываются при фиксированном значении n. Следовательно, $x_n z$ и их предел xz обладают желаемой измеримостью по Борелю. Теперь можно применить это к $x\xi_1, x\xi_2, \ldots$, и так как их предел ξ существует почти наверное в силу теоремы Банаха о неподвижной точке (см. замечания 7.3.$_1$), то утверждение доказано.

Так, мы приходим к решению путем итераций стохастического оператора. Что будет, если образовать произведение независимых стохастических операторов, аналогичных типу, описанному выше? Этот вопрос возникает в связи со многими прикладными задачами, но следующую теорему можно воспринимать только как первую попытку подойти к решению и надо ожидать более сильных результатов.

Теорема 7.3.2. *Пусть Z — банахово пространство, и пусть множество X состоит из операторов, отображающих Z в Z и таких, что элементы x удовлетворяют соотношению*

$$\| (xz_1 - xz_2) \| \leqslant \varrho(x) \| z_1 - z_2 \|,$$

где $\varrho(x)$ — случайная величина с математическим ожиданием ϱ, меньшим единицы, и $\mathbf{E} \| x_1 x_2 \ldots x_n \xi \|$ равномерно ограничено для любого фиксированного ξ. Возьмем независимые случайные элементы $x_1, x_2, \ldots$ из X и образуем случайные элементы $\xi_n = x_1 x_2 \ldots x_n \xi$. При $n \to \infty$ случайные элементы ξ_n сходятся в среднем L_1 к случайному элементу ζ, не зависящему от начального элемента ξ.

Доказательство. Образуем для произвольного n и $h > 0$ разность $\xi_{n+h} - \xi_n = d_{n,h}$. Она имеет норму

$$\| d_{n,h} \| = \| x_1 x_2 \ldots x_n x_{n+1} \ldots x_{n+h} \xi - x_1 x_2 \ldots x_n \xi \| \leqslant$$
$$\leqslant \varrho(x_1) \varrho(x_2) \ldots \varrho(x_n) \cdot \| x_{n+1} x_{n+2} \ldots x_{n+h} \xi - \xi \|,$$

откуда следует сходимость. Для двух различных ξ_1, $\xi_2 \in Z$ имеем

$$\mathbf{E}\,\|d_n\| = \mathbf{E}\,\|x_1x_2\ldots x_n\xi_1 - x_1x_2\ldots x_n\xi_2\| \leqslant \leqslant \varrho^n \|\xi_1-\xi_2\| \to 0,$$

так что

$$\mathbf{E}\,\|d_n\| \leqslant \varrho^n \cdot C \to 0 \quad (C \text{ — константа})$$

при $n\to\infty$, и предел ζ не зависит от ξ.

Если задана последовательность описанных выше случайных операторов $\ldots,\ x_{-1},\ x_0,\ x_1,\ \ldots$, то можно образовать последовательность случайных элементов

$$\xi_n = \lim_{\nu\to\infty} x_n x_{n-1} x_{n-2} \ldots x_{n-\nu}\xi,$$

где ξ — некоторый произвольный начальный элемент. Тогда новая последовательность ξ_n, принимающая значения из Z, имеет стационарное распределение вероятностей, и мы можем говорить, что это — стационарное решение *случайного уравнения*

$$\xi_{n+1} = x_n\xi_n, \quad n = \ldots,\ -1,\ 0,\ 1,\ \ldots\ .$$

Можно также отметить, что не существенно то, что Z есть банахово пространство: это предположение может быть ослаблено, и можно работать с более общими метрическими пространствами.

Вполне возможно, что последняя теорема может быть обобщена и уточнена. Но все-таки какое-то условие, похожее на свойство сжатости, видимо, необходимо для того, чтобы величины ξ_n сходились. Некоторое внимание было уделено связанной с этим проблеме: выяснить, когда все стохастические преобразования являются сохраняющими меру. Это приводит нас к так называемой *случайной эргодической теореме*, утверждающей не то, что ξ_n сходятся, а то, что они обладают некоторым средним поведением.

Рассмотрим вероятностное пространство $(\Omega, \mathscr{B}_\Omega, P)$, и пусть каждому $\omega\in\Omega$ соответствует некоторое преобразование x_ω другого вероятностного пространства $(Z, \mathscr{B}_Z, m)$ в себя. Будем предполагать, что x_ω — сохраняющее меру преобразование, так что $x_\omega(B)$ и $x_\omega^{-1}(B)\in\mathscr{B}_Z$ для любого $B\in\mathscr{B}_Z$ и $m[x_\omega(B)] = m(B)$. Семейство $\{x_\omega;\ \omega\in\Omega\}$

должно быть измеримым в том смысле, что $\{(\omega, z) \mid x_\omega z \in B\} \in \mathscr{B}_\Omega \times \mathscr{B}_Z$. Введем теперь обычным образом новое вероятностное пространство $(\Theta, \mathscr{B}_\Theta, P_\Theta)$, состоящее из последовательностей $\theta = (\ldots, \omega_{-1}, \omega_0, \omega_1, \ldots)$, где каждая координата ω_n имеет распределение вероятностей P и все координаты независимы. На Θ имеем преобразование переноса φ, определенное так, что $\varphi\theta$ — новая последовательность, n-я координата которой равна θ_{n+1}. Ясно, что φ есть сохраняющее меру преобразование.

Пусть теперь $f(z)$ — m-интегрируемая действительная функция, определенная на Z. Образуем среднее

$$S_n = \frac{1}{n} \sum_{\nu=0}^{n-1} f\left[x_{\omega_{\nu-1}} x_{\omega_{\nu-2}} \ldots x_{\omega_0} z\right].$$

Уместно рассмотреть произведение пространств с мерами $Z \times \Theta$ и ввести преобразование

$$x(z, \theta) = [x_{\omega_0} z, \varphi\theta].$$

Это новое преобразование является сохраняющим меру, и можно применить обычную индивидуальную эргодическую теорему к пространству $Z \times \Theta$. Так как ν-я итерация преобразования $x(z, \theta)$ определяется соотношением

$$x^\nu(z, \theta) = [x_{\omega_{\nu-1}} \ldots x_{\omega_1} x_{\omega_0} z, \varphi^\nu \theta],$$

то имеем

$$S_n = \frac{1}{n} \sum_{\nu=0}^{n-1} f\ [x\text{-координата } x^\nu(z, \theta)],$$

и мы знаем, что почти наверное на пространстве $Z \times \Theta$ среднее S_n сходится к некоторому интегрируемому пределу. Применяя теорему Фубини, получаем следующий результат (см. замечания 7.3.$_2$).

Теорема 7.3.3. (Случайная эргодическая теорема.) *При сделанных предположениях существует событие $N \subset \Theta$ вероятности нуль, $P(N) = 0$, такое, что если*

$$\theta = (\ldots, \omega_{-1}, \omega_0, \omega_1, \ldots) \overline{\in} N,$$

то

$$\lim_{n\to\infty}\frac{1}{n}\sum_{\nu=0}^{n-1} f\,[x_{\omega_{\nu-1}}x_{\omega_{\nu-2}}\ldots x_{\omega_0}z]=\bar{f}(z)$$

существует для почти всех $z \in Z$ и $\bar{f}(z) \in L_1(Z)$.

Вообще говоря, нельзя утверждать, что предельная функция $\bar{f}(z)$ почти наверное является постоянной.

7.4. Более специальные структуры

Если наложить на банахову алгебру X дополнительные требования, то относительно распределений вероятностей на ней можно, конечно, сказать несколько больше. В настоящее время такая специализация не привела к интересным результатам, может быть, потому, что, идя прямым путем, мы приходим к хорошо известным структурам.

Можно было бы потребовать, чтобы банахова алгебра, кроме операций сложения, умножения и умножения на скаляр, обладала бы четвертой операцией — делением. Более точно, если предположить, что X — *полная нормированная алгебра с делением*, так что x^{-1} существует и является непрерывной функцией от x при $x \neq 0$, то X изоморфно полю комплексных чисел (см. замечания 7.4.$_1$). Но вряд ли можно ожидать, что распределения вероятностей на комплексной плоскости представят увлекательный предмет для исследования. Возьмем, например, предельные теоремы. Если записывать комплексные числа в полярных координатах $x = re^{i\varphi}$, то

$$x_1x_2\ldots x_n = r_1r_2\ldots r_n \exp\,[i\,(\varphi_1+\varphi_2+\ldots+\varphi_n)].$$

Предполагается, что множители x_ν независимы и имеют одинаковые распределения. Пусть F обозначает совместное распределение $\log r = \varrho$ и φ. Нужно рассмотреть сложение на коммутативной группе $R^1 \times T^1$, состоящей из элементов (ϱ, φ). Характеры имеют, конечно, вид $\exp(iu\varrho + iv\varphi)$, где u — действительное, а v — целое число. Предположим, что φ имеет нерешетчатое распределение, так что $|\mathbf{E}\exp(iv\varphi)| < 1$ для $v \neq 0$. Введем преобразование

Фурье

$$\hat{F}(u, v)=\int_{-\infty}^{\infty}\int_{0}^{2\pi}\exp(iu\varrho+iv\varphi)\,dF(\varrho, \varphi)$$

и предположим, что моменты

$$m=\mathbf{E}\varrho,$$
$$\sigma^2=\mathbf{E}(\varrho-m)^2$$

конечны. Тогда величина

$$(P_n, \Phi_n)=\left(\frac{\varrho_1+\varrho_2+\ldots+\varrho_n-nm}{\sqrt{n}},\ \varphi_1+\varphi_2+\ldots+\varphi_n\right)$$

имеет распределение вероятностей с преобразованием Фурье

$$\left[\hat{F}\left(\frac{u}{\sigma\sqrt{n}}, v\right)\exp\left(-\frac{imu}{\sqrt{n}\,\sigma}\right)\right]^n.$$

Когда n стремится к бесконечности,

$$\hat{F}\left(\frac{u}{\sigma\sqrt{n}}, v\right)\to\hat{F}(0, v)=\mathbf{E}\exp(iv\varphi),$$

что по абсолютной величине меньше единицы при $v\neq 0$. Следовательно,

$$\left[\hat{F}\left(\frac{u}{\sigma\sqrt{n}}, v\right)\right]^n\to 0 \quad\text{при}\quad v\neq 0.$$

Для $v=0$ получаем, как в доказательстве простейшего варианта центральной предельной теоремы, что

$$\left[\hat{F}\left(\frac{u}{\sigma\sqrt{n}}, 0\right)\exp\left(-\frac{imu}{\sqrt{n}}\right)\right]^n\to\exp\left(-\frac{1}{2}u^2\right).$$

Это показывает, что при $n\to\infty$ асимптотическое распределение (P_n, Φ_n) нормально для P_n, равномерно для Φ_n, и эти две координаты независимы. В случае, когда φ имеет решетчатое распределение, это рассуждение может быть изменено. Этот случай имеет место, например, когда X — действительное поле, $\varphi=0$ или $\varphi=\pi$.

Другой случай, близкий к рассмотренному, получим, когда банахова алгебра X является *коммутативной* или *полупростой*. Из теории Гельфанда известно, что в последнем случае X изоморфно семейству непрерывных функций $c(t)$ на некотором компактном множестве T (см. замечания 7.4.$_2$). И снова ситуация представляет мало интереса. Для того чтобы изучать предельные теоремы, мы будем рассматривать произведения $c_1(t)\,c_2(t)\,\ldots\,c_n(t)$ и суммы $c_1(t)+c_2(t)+\ldots+c_n(t)$, приводящие соответственно к логонормальному и равномерному (может быть, на некоторой подгруппе) распределениям или к нормальному распределению. Это имеет место, когда рассматривается значение произведения или суммы для каждого t отдельно, но можно также определить совместное асимптотическое распределение для любого конечного числа значений t или на σ-алегбре, порожденной соответствующими цилиндрическими множествами.

7.5. Примеры

7.5.1. Важный случай стохастической алгебры представляют собой *случайные* $k\times k$-матрицы. Эта алгебра обладает достаточно богатой структурой для того, чтобы на ней могла быть развита полезная и содержательная теория вероятностей. Рассмотрим сейчас некоторые ее аспекты. Если записать $x\in X$ в виде

$$x=\{x_{ij};\ i,\ j=1,\ 2,\ \ldots,\ k\}$$

и ввести норму

$$\|x\|=\max_i\sum_j|x_{ij}|,$$

то X будет банаховой алгеброй. Если предположить, что случайные величины x_{ij} имеют средние значения m_{ij}, то, конечно, x будет иметь среднее значение

$$m=\{m_{ij};\ i,\ j=1,\ 2,\ \ldots,\ k\}.$$

Если $x_1,\ x_2,\ \ldots$ — независимые стохастические матрицы в X со средними значениями $m_1,\ m_2,\ \ldots\in X$, то, очевидно, произведение $x_1x_2\,\ldots\,x_n$ имеет среднее значение $m_1m_2\,\ldots\,m_n$. Это утверждение справедливо для произвольной банаховой алгебры. Действительно, если y и η — два стохастически независимых элемента в некоторой банаховой алгебре X со средними значениями m и μ, то для $z=y\eta$ случайная величина $\|z\|\leqslant\|y\|\cdot\|\eta\|$ интегрируема (z, конечно, измерима по Борелю). Тогда $\mathbf{E}z=\nu$ существует, и для любого ограниченного линейного функционала $x^*\in X^*$ мы должны иметь

$$x^*(\nu)=\mathbf{E}x^*(z)=\int\int x^*(y\eta)\,P_1(dy)\,P_2(d\eta).$$

Но $x^*(z)=x^*(y\eta)=x^*_y(\eta)$ — ограниченный линейный функционал x^*_y от η при фиксированном y. Тогда теорема Фубини дает нам

$$x^*(\nu)=\int x^*_y(\mu)\,P_1(dy)=\int x^*(y\mu)\,P_1(dy)=x^*(m\mu),$$

так что $\nu=\mathbf{E}y\eta=m\mu=\mathbf{E}y\cdot\mathbf{E}\eta$, что и утверждалось.

Получаем, в частности, $\mathbf{E}x_1x_2\ldots x_n=m^n$, если все множители x_ν имеют одно и то же среднее значение m. В матричном случае известно, что поведение m^n при больших n связано со свойствами наибольших собственных значений m (можно использовать, например, каноническую форму Жордана). В общем случае оно более сложным образом зависит от спектральных свойств элемента m.

Рассмотрим теперь $y_n=\{y^{(n)}_{ij};\ i,\ j=1,\ 2,\ \ldots,\ k\}$, где $y_n=x_1x_2\ldots x_n$. Тогда, предполагая моменты второго порядка существующими, получаем

$$\mathbf{E}y^{(n)}_{ij}y^{(n)}_{\alpha\beta}=\sum\mathbf{E}x^{(1)}_{ii_1}x^{(1)}_{\alpha\alpha_1}\mathbf{E}x^{(2)}_{i_1i_2}x^{(2)}_{\alpha_1\alpha_2}\ldots\mathbf{E}x^{(n)}_{i_{n-1}j}x^{(n)}_{\alpha_{n-1}\beta},$$

где суммирование распространяется на $\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_{n-1}$, $i_1,\ i_2,\ \ldots,\ i_{n-1}$. Вводя ковариационные квадратные матрицы порядка k^2

$$\begin{cases} C=\{C_{ij,\,\alpha\beta};\ i,\ j,\ \alpha,\ \beta=1,\ 2,\ \ldots,\ k\},\\ \qquad C_{ij,\,\alpha\beta}=\mathbf{E}x^{(1)}_{ij}x^{(1)}_{\alpha\beta},\end{cases}$$

получаем

$$\{\mathbf{E}y^{(n)}_{ij}y^{(n)}_{\alpha\beta}\}=C^n.$$

Снова можно сделать некоторые асимптотические утверждения. Можно перейти также к моментам более высокого порядка.

Хотя подобная выкладка и не лишена интереса, она мало помогает в том, что касается предельных теорем.

Это ясно уже на действительной прямой. В самом деле, если $X = R^1$, то нам известно асимптотическое поведение $y_n = x_1 x_2 \ldots x_n$, которое может быть выражено в терминах логонормального распределения. Если $a = \mathbf{E} \log x$ и $b = \mathbf{D} (\log x)$ существуют, то $(\log y_n - na)(nb)^{-1/2}$ асимптотически нормально. Тогда можно ожидать, что y_n имеет порядок e^{na}, а $\mathbf{E} y_n = (\mathbf{E} x)^n$; первое выражение соответствует среднему геометрическому, а второе — среднему арифметическому, и, вообще говоря, они существенно различны.

Из разд. 7.1 известно, что при изучении некоторых предельных задач мы можем рассматривать соответствующие мультипликативные однородные процессы, зависящие от непрерывного параметра времени t. Рассмотрим два таких процесса в матричном случае.

Пусть $y(t)$, $t \geqslant 0$, есть аддитивный однородный случайный процесс, значениями которого являются $k \times k$-матрицы и который ведет себя подобно пуассоновскому процессу, так что

$$y(t+h) = \begin{cases} y(t) + hA & \text{с вероятностью } 1 - \lambda h + o(h), \\ y(t) + hA + B & \text{с вероятностью } \lambda h + o(h). \end{cases}$$

Более точно, мы исходим из обычного пуассоновского процесса с моментами времени t_ν появления событий, наступающих с плотностью λ. Между двумя такими t_ν процесс изменяется линейно, пропорционально постоянной матрице A. В момент t_ν процесс y «подскакивает» на матрицу B, также постоянную. Это определяет распределение вероятностей данного процесса. Он удовлетворяет условию C_1 (разд. 7.1), так как

$$\sum \mathbf{E} \| y(t_\nu) - y(t_{\nu-1}) \| \leqslant$$

$$\leqslant \sum (t_\nu - t_{\nu-1}) [\| A \| + \lambda \| B \| + o(1)] = O(1).$$

Следовательно, соответствующий мультипликативный однородный процесс $x(t)$ определен. Если A и B коммутируют, то $x(t)$ есть просто $x(t) = \exp[tA + n(t)B]$, где $n(t)$ — случайная величина, распределенная по закону Пуассона. Если же $AB \neq BA$, то процесс x более сложен. Во всяком случае, мы можем вычислить достаточно просто его средние значения и вторые моменты. Полагая

$$M(t)=\{m_{ij}(t);\ i,\ j=1,\ 2,\ \ldots,\ k\},\quad m_{ij}(t)=\mathbf{E}x_{ij}(t),$$

получаем

$$M(t+h)=\mathbf{E}x(t+h)=M(t)[I+hA+h\lambda B+o(h)],$$

так что

$$M'(t)=M(t)(A+\lambda B)$$

и

$$M(t)=\exp[t(A+\lambda B)].$$

Чтобы определить моменты второго порядка, введем $k^2\times k^2$-матрицу

$$C(t)=\{c_{ij,\,\alpha\beta}(t);\ i,\ j,\ \alpha,\ \beta=1,\ 2,\ \ldots,\ k\}=\\ =\{\operatorname{cov}[x_{ij}(t),\ x_{\alpha\beta}(t)]\},$$

где индексы упорядочиваются, скажем, в словарном порядке. Если мы обозначим через $R\times S$ кронекеровское произведение двух матриц R и S, то получим

$$C(t)=\{\mathbf{E}x_{ij}(t)x_{\alpha\beta}(t)-m_{ij}(t)m_{\alpha\beta}(t)\}=\\ =\mathbf{E}x(t)\times x(t)-M(t)\times M(t).$$

Но матрицы $S(t)=\mathbf{E}x(t)\times x(t)$ удовлетворяют соотношению

$$S(t+h)=\mathbf{E}x(t+h)\times x(t+h)=\\ =\mathbf{E}[x(t)+x(t)\Delta y]\times[x(t)+x(t)\Delta y]+$$

$+$ члены более высокого порядка малости, где положено $\Delta y=y(t+h)-y(t)$. Следовательно,

$$S(t+h)=S(t)+h\mathbf{E}x(t)\times x(t)[A+\lambda B]+\\ +h\mathbf{E}x(t)[A+\lambda B]\times x(t)+\mathbf{E}x(t)\Delta y\times x(t)\Delta y+\ldots,$$

так что

$$S'(t)=S(t)(I\times[A+\lambda B])+S(t)([A+\lambda B]\times I)+\lambda S(t)B\times B$$

и

$$S(t)=\exp\{t[I\times(A+\lambda B)+(A+\lambda B)\times I+\lambda B\times B]\}.$$

Можно установить асимптотическое поведение $M(t)$ и $S(t)$ при больших t и распространить полученные результаты на моменты более высокого порядка.

Будем теперь исходить из другого процесса y. Предположим, что $y(t)$ имеет нормальные распределения с независимыми приращениями и средними значениями, равными нулю. Вторые моменты пропорциональны t, и мы пишем

$$\mathbf{E} y_{ij}(t)\, y_{\alpha\beta}(t) = t b_{ij,\,\alpha\beta}.$$

Подходящее определение нормы задается соотношением

$$\| U \|^2 = (U, U),$$

где

$$(U, V) = \sum_{i,\,j} u_{ij} v_{ij},$$

а U и V — действительные $k \times k$-матрицы с элементами соответственно u_{ij} и v_{ij}. Чтобы проверить условие C_2, нужно только рассмотреть $y(R)$, где R — параллелепипед в S_r (см. разд. 7.1.). Но

$$\| y(R) \|^2 = \sum y^{(1)}_{i_1 i_2} y^{(2)}_{i_2 i_3} \ldots y^{(r)}_{i_r i_{r+1}} y^{(1)}_{i_1 j_1} y^{(2)}_{j_1 j_2} \ldots y^{(r)}_{j_r i_{r+1}},$$

где сумма распространяется на все значения индексов между 1 и k и $y^{(\nu)}_{ij}$ есть (ij)-й элемент матрицы, представляющей приращение $y(t)$ вдоль ν-го ребра длины Δ_ν параллелепипеда R. Отсюда вытекает, что

$$\mathbf{E} \| y(R) \|^2 = \sum \mathbf{E} y^{(1)}_{i_1 i_2} y^{(1)}_{i_1 j_1} \cdot \mathbf{E} y^{(2)}_{i_2 i_3} y^{(2)}_{j_1 j_2} \ldots \mathbf{E} y^{(r)}_{i_r i_{r+1}} y^{(r)}_{j_r i_{r+1}} =$$
$$= O(k^{2r+1} b^r \Delta_1 \Delta_2 \ldots \Delta_r) = O(C^r \cdot m(R)),$$

где $b = \max | b_{ij\alpha\beta} |$, C — константа, и это показывает, что наше условие выполнено, так что соответствующий мультипликативный однородный процесс $x(t)$ вполне определен.

Инфинитезимальные ковариации $x(t)$ имеют вид

$$\lim_{h \to 0} 1/h \ \{\text{условная ковариация } [x_{ij}(t+h) -$$
$$- x_{ij}(t),\ x_{\alpha\beta}(t+h) - x_{\alpha\beta}(t)]\} =$$
$$= \sum_{l,\,\lambda=1}^{k} x_{il}(t)\, x_{\alpha\lambda}(t)\, b_{lj,\,\lambda\beta} = q_{ij,\,\alpha\beta}[x(t)],$$

где условная ковариация берется при фиксированном значении $x(t)$. Величины $q_{ij,\,\alpha\beta}(x)$ являются квадратичными формами элементов матрицы x, и $k^2 \times k^2$-матрица $q=\{q_{ij,\alpha\beta}\}$ может быть записана в виде $[x(t) \times x(t)]\, B$. Тогда уравнение Фоккера — Планка имеет вид

$$\frac{\partial p}{\partial t}=\frac{1}{2}\sum\frac{\partial^2}{\partial x_{ij}\,\partial x_{\alpha\beta}}[q_{ij,\,\alpha\beta}(x)\,p],$$

и для его решения должны быть выбраны соответствующие начальные условия. В нашем случае следует принять $\mathbf{P}\,\{x(0)=I\}=1$.

Если средние значения процесса y не равны нулю, а

$$\mathbf{E}y(t)=tA, \quad A=\{a_{ij}\},$$

то приведенное выше уравнение будет иметь вид

$$\frac{\partial p}{\partial t}=\frac{1}{2}\sum\frac{\partial^2}{\partial x_{ij}\,\partial x_{\alpha\beta}}[q_{ij,\,\alpha\beta}(x)\,p]-\sum\frac{\partial}{\partial x_{ij}}[l_{ij}(x)\,p],$$

где $L=\{l_{ij}(x)\}$ и $L=x(t)\,A$, так что $l_{ij}(x)$ являются линейными формами переменных x_{ij}.

Чтобы найти ожидаемые значения процесса x, возьмем частные производные

$$\frac{\partial}{\partial t}\mathbf{E}x_{ij}(t)=\frac{\partial}{\partial t}m_{ij}(t)=\mathbf{E}l_{ij}(x),$$

так что

$$\frac{\partial m(t)}{\partial t}=\mathbf{E}x(t)\,A=m(t)\,A; \quad m(t)=\exp(tA).$$

Аналогично для моментов второго порядка получаем

$$\frac{\partial}{\partial t}c_{ij,\,\alpha\beta}(t)=\mathbf{E}q_{ij,\,\alpha\beta}[x(t)]+\mathbf{E}l_{ij}(x)\,x_{\alpha\beta}(t)+\mathbf{E}l_{\alpha\beta}(x)\,x_{ij}(t),$$

или

$$\frac{\partial c(t)}{\partial t}=c(t)\,B+c(t)\,[A\times I+I\times A],$$

так что

$$c(t)=\exp[t(B+A\times I+I\times A)].$$

Сравните это с аналогичными выражениями, полученными выше в этом разделе.

Изучавшийся выше процесс x интересен при рассмотрении многих предельных задач. Поэтому существенно важно найти методы, позволяющие получать решения этих уравнений Фоккера — Планка. Конечно, можно было бы решать непосредственно эти уравнения численными методами, но это имеет ограниченный практический интерес, за исключением частных случаев, так как большое число параметров делает табулирование затруднительным. В самом деле, мы имеем k^2 параметров для инфинитезимальных средних значений и $k^2(k^2+1)/2$ параметров для инфинитезимальных ковариаций, всего $N = k^2(k^2+3)/2$ параметров. При $k=2$ имеем $N=14$, и уже для $k=3$ получаем устрашающее число: $N=54$ параметрам. Это делает почти необходимым использование аналитического или аппроксимационного подхода, скажем, выражения решения через табулированные функции, или применение подходящего разложения в ряд, или применение какого-либо интегрального представления. Очевидно, что эта проблема нуждается в дальнейшем изучении.

Прежде чем перейти к другим вопросам, сделаем несколько качественных замечаний. Для простоты рассмотрим стохастические квадратные матрицы M_h второго порядка с равными нулю средними значениями, действующие на векторы x из R^2, $\Delta x = x(t+h) - x(t) \cong M_h x(t)$. Распределение $\Delta x \in R^2$ определяется инфинитезимальными ковариациями

$$q_{ij}(x) = \operatorname{cov}[(Mx(t))_i, \quad (Mx(t))_j],$$

вычисленными для некоторого $x(t) = x$, где M — случайная квадратная матрица второго порядка. Если матрица $Q(x) = \{q_{ij}(x);\ i, j = 1, 2\}$ невырождена в x, то диффузия имеет место во всех направлениях из точки x. Если мы потребуем, чтобы распределение вероятностей ограничивалось некоторым подмножеством пространства R^2, то матрица $Q(x)$ будет вырожденной при некоторых значениях x, скажем, на некотором подмножестве S. Но M имеет четыре элемента, так что она может быть представлена в виде

$$M = \sum_{\nu=1}^{r} \xi_\nu A_\nu,$$

где A_ν — неслучайные квадратные матрицы второго порядка, ξ_ν — некоррелированные случайные величины с единичной дисперсией, а $r \leqslant 4$. Тогда

$$q_{ij}(x) = \sum_{\nu=1}^{r} (A_\nu x)_i (A_\nu x)_j,$$

так что, если z — произвольный вектор-столбец, то

$$z^* Q(x) z = \sum_{\nu=1}^{r} (A_\nu x, z)^2.$$

Если эта квадратическая форма обращается в нуль для некоторого нетривиального вектора z, то диффузия в направлении вектора z не имеет места. Пусть S — множество значений x из R^2, для которых $Q(x)$ является вырожденной:

$$S = \{x \mid \det Q(x) = 0\},$$

так что в S имеем

$$0 = \det Q(x) = \sum_{\nu,\,\mu} \begin{vmatrix} (A_\nu x)_1 (A_\mu x)_1 & (A_\nu x)_1 (A_\mu x)_2 \\ (A_\nu x)_2 (A_\mu x)_1 & (A_\nu x)_2 (A_\mu x)_2 \end{vmatrix} =$$

$$= \sum_{\nu<\mu} \begin{vmatrix} (A_\nu x)_1 & (A_\mu x)_1 \\ (A_\nu x)_2 & (A_\mu x)_2 \end{vmatrix}^2,$$

откуда вытекает, что

$$\begin{vmatrix} (A_\nu x)_1 & (A_\mu x)_1 \\ (A_\nu x)_2 & (A_\mu x)_2 \end{vmatrix} = 0$$

для всех ν и μ. Но это — множество однородных уравнений второго порядка относительно x, так что S может быть изолированной точкой $x = 0$, одной или двумя прямыми линиями или всей плоскостью.

Рассмотрим сначала $r = 1$. Тогда $S = R^2$, и квадратическая форма $z^* Q(x) z$ равна нулю, если $(Ax, z) = 0$. Диффузия имеет место только в направлении, параллельном вектору Ax, так что мы приходим к дифференциальному уравнению для особых кривых

$$\frac{dx_2}{dx_1} = \frac{(Ax)_2}{(Ax)_1},$$

рассматриваемому во всех элементарных учебниках по дифференциальным уравнениям. Его решения образуют семейства кривых C_α, состоящие из прямых линий, эллипсов, спиралей, гипербол и парабол различных типов. Если начальное распределение сосредоточено на C_α, то распределенная масса никогда не покидает этой кривой, но может только диффундировать вдоль нее. Если A вырождена, то мы получаем еще более вырожденную ситуацию, так как если вектор x аннулирует A_1, $A_1x = 0$, то масса, находящаяся в x, остается там неопределенно долго.

Рассмотрим теперь $r = 2$. Тогда S состоит из точек, в которых

$$\det A(x) = \begin{vmatrix} (A_1x)_1 & (A_2x)_1 \\ (A_1x)_2 & (A_2x)_2 \end{vmatrix} = 0,$$

откуда вытекает, что $(A_1 - \lambda A_2)\,x = 0$ для некоторого скаляра λ. Для нетривиального вектора x мы должны иметь $\det(A_1 - \lambda A_2) = 0$, так что возможны только два различных значения λ. Исключая случай, когда A_1 и A_2 кратны друг другу ($A_1 = \lambda A_2$), что соответствует случаю $r = 1$, мы видим, что уравнение $(A_1 - \lambda A_2)\,x = 0$ имеет решения $x = \alpha u$, где $u \in R^2$ и α — произвольный скаляр. Когда такая прямая линия является особой в описанном смысле? Если диффузия может происходить только в направлении u, то мы должны иметь

$$0 = (A_1x_1n)^2 + (A_2x_1n)^2 = \alpha^2(A_1u_1n)^2 + \alpha^2(A_2u_1n)^2,$$

где n — вектор, ортогональный u, так что

$$A_1u = \mu_1u,$$
$$A_2u = \mu_2u.$$

Это означает, что A_1 и A_2 должны иметь один или два общих правых собственных вектора, и мы получим одну или две особые прямые линии.

Для $r = 3$ или 4 мы получаем аналогичные ответы. Особые кривые имеют место только тогда, когда у матриц A_ν есть общие собственные векторы; в противном случае распределенная масса заполняет всю плоскость ($x = 0$ — всегда, конечно, особая точка).

Пусть $E \subset R^2$ есть область, конечная часть которой ограничена особыми кривыми. Если начальное распределение сосредоточено на E, то интуитивно ясно, что никакая часть распределенной массы не покинет E, $P_t(E) \equiv 1$.

Для того чтобы показать это, предположим для простоты, что E ограничена двумя лучами L_1 и L_2, выходящими из начала координат. Удобно использовать полярные координаты ϱ и θ. Пусть лучам L_1 и L_2 соответствуют полярные углы θ_1 и θ_2. Легко выписать уравнение Фоккера — Планка

$$\frac{\partial}{\partial t} p(\theta, \varrho; t) = \frac{1}{2} \frac{\partial^2}{\partial \theta^2} [k_1(\theta) p] + \frac{\partial^2}{\partial \theta \, \partial \varrho} [\varrho k_2(\theta) p] + \frac{1}{2} \frac{\partial^2}{\partial \varrho^2} [\varrho^2 k_3(\theta) p],$$

где $k_i(\theta)$ — тригонометрические полиномы второго порядка относительно переменной θ, а $k_1(\theta)$ а $k_2(\theta)$ обращаются в нуль для $\theta = \theta_1$ и $\theta = \theta_2$. Теперь мы можем проинтегрировать по E,

$$P_t(E) = \int_{\theta_1}^{\theta_2} \int_0^{\infty} \varrho p(\theta, \varrho; t) \, d\theta \, d\varrho,$$

и после нескольких интеграций получим $\frac{\partial}{\partial t} P_t(E) = 0$, что и утверждалось. Мы использовали тот факт, что $k_1(\theta)$ имеет нули второго порядка, а $k_2(\theta)$ — нули первого порядка при $\theta = \theta_1$ и $\theta = \theta_2$.

Перейдем теперь к другому аспекту той же проблемы — поведению угла $\theta(t)$ для больших значений t. Предположим, что особых линий нет, $k_1(\theta) > 0$. Мы не будем больше игнорировать инфинитезимальное среднее значение величины θ и не будем идентифицировать углы, различающиеся на кратное 2π. Вместо этого мы будем следить за тем, как непрерывно меняется $\theta(t)$ с изменением t (см. также 4.4.1). Заметим, что $\theta(t)$, так же как и $(\theta(t), \varrho(t))$, образует марковский процесс, так как $\Delta x = M_h x$ есть однородная функция от переменного x. Уравнение Фоккера — Планка для плотности $\pi(\theta; t)$ этого угла $\theta(t)$ может быть записано в виде

$$\frac{\partial \pi(\theta; t)}{\partial t} = \frac{1}{2} \frac{\partial^2}{\partial \theta^2} [k_1(\theta) \pi] - \frac{\partial}{\partial \theta} [k_4(\theta) \pi],$$

где $k_4(\theta)$ — тригонометрический полином первого порядка. Оно может быть выведено также из приведенного ранее уравнения Фоккера — Планка (с дополнительными членами, содержащими средние значения) интегрированием по переменной ϱ. Имеем

$$\frac{d}{dt}\,\mathbf{E}\theta(t)\int\limits_{-\infty}^{\infty}\theta\frac{\partial\pi(\theta;\,t)}{\partial t}\,d\theta=$$

$$=\int\limits_{-\infty}^{\infty}\pi(\theta;\,t)\,k_4(\theta)\,d\theta=\int\limits_{0}^{2\pi}p(\theta;\,t)\,k_4(\theta)\,d\theta,$$

где последнее соотношение вытекает из того, что $k_4(\theta)$ периодическая функция с периодом 2π. Предположим, что мы исходим из процесса с существующей стационарной плотностью $p(\theta)$. Для вычисления $p(\theta)$ нужно решить уравнение

$$\frac{1}{2}\frac{\partial^2}{\partial\theta^2}[k_1(\theta)\,p]-\frac{\partial}{\partial\theta}[k_4(\theta)\,p]=0,$$

или

$$\frac{1}{2}\frac{\partial}{\partial\theta}[k_1(\theta)\,p]-k_4(\theta)\,p=a,$$

имеющее решения

$$p(\theta)=\frac{1}{k_1(\theta)}\left[b+2a\int\limits_0^{\theta}\exp\left(-2\int\limits_0^{u}\frac{k_4(v)}{k_1(v)}\,dv\right)du\right]\times$$

$$\times\exp\left(2\int\limits_0^{\theta}\frac{k_4(u)}{k_1(u)}\,du\right),$$

где константы a и b должны быть выбраны так, чтобы $p(\theta)$ являлось плотностью, удовлетворяющей граничному условию $p(2\pi)=p(0)$. Но

$$p(0)=\frac{b}{k_1(0)},$$

$$p(2\pi)=\frac{b+2aI_1}{k_1(0)}\exp I_2,$$

где

$$I_1=\int\limits_0^{2\pi}\exp\left(-2\int\limits_0^{u}\frac{k_4(v)}{k_1(v)}\,dv\right)du>0,$$

$$I_2=2\int\limits_0^{2\pi}\frac{k_4(u)}{k_1(u)}\,du.$$

Это приводит нас к условиям

$$b = (b + 2aI_1) \exp I_2,$$

так что

$$a = \frac{1 - \exp I_2}{2I_1 \exp I_2} b.$$

Выбирая b как некоторое положительное число, можно сделать $p(\theta)$ положительной функцией, так как $b + 2aI_1 = b \exp(-I_2) > 0$, и такой, что интеграл от нее по интервалу $(0, 2\pi)$ равен единице. Комбинируя все это, видим, что

$$\frac{d}{dt} \mathbf{E}\theta(t) = \int_0^{2\pi} p(\theta) k_4(\theta) d\theta = -2\pi a,$$

так что

$$\mathbf{E} \frac{\theta(t) - \theta(0)}{t} = -2\pi a.$$

Иными словами, точка вращается вокруг начала координат со средней угловой скоростью a оборотов в единицу времени в отрицательном направлении.

Этот пример может иметь некоторый общий интерес, если его рассмотреть с точки зрения разд. 4.5.1.

7.5.2. Предположим теперь, что все возможные элементы матриц x положительны и удовлетворяют неравенству вида

$$\frac{\max x_{ij}}{\min x_{ij}} \leqslant C < \infty.$$

Для произведения $y^n = \{y_{ij}^n\} = x^1 x^2 \ldots x^n$ имеем

$$\min_{i, j} y_{ij}^n \leqslant \| y^n \| \leqslant k \max_{i, j} y_{ij}^n,$$

где используется та же норма, что и в начале разд. 7.5.1. Но

$$\frac{y_{ij}^n}{y_{\alpha\beta}^n} = \frac{\sum\limits_{\nu, \mu} x_{i\nu}^n z_{\nu\mu} x_{\mu j}^1}{\sum\limits_{\nu, \mu} x_{\alpha\nu}^n z_{\nu\mu} x_{\mu\alpha}^1} \leqslant C^2,$$

где $\{z_{\nu\mu}\} = x^2 x^3 \ldots x^{n-1}$, так что

$$\max_{i, j} y_{ij}^n \leqslant C^2 \min_{i, j} y_{ij}^n.$$

Следовательно,

$$\frac{1}{n}\log\min_{i,\,j} y_{ij}^{n} \leqslant \frac{1}{n}\log\|y^{n}\| \leqslant \frac{1}{n}\log(kC^{2}) + \frac{1}{n}\log\min_{i,\,j} y_{ij}^{(n)}.$$

Средняя часть неравенства стремится почти наверное к пределу r (если r конечно; см. замечания 7.1 и разд. 7.1). Отсюда и из приведенного выше неравенства вытекает, что

$$\frac{1}{n}\log y_{ij}^{n} \to r$$

почти наверное (если r конечно) для всех индексов i и j.

Было показано (см. замечания 7.5.2.), что если $\mathbf{E}\,|\log x_{11}|^{2+\delta} < \infty$ для некоторого положительного δ, то существуют действительные постоянные a и b, такие, что

$$\lim_{n\to\infty}\mathbf{P}\left\{\frac{\log y_{ij}^{n} - na}{\sqrt{nb}} \leqslant x\right\} = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x} e^{-z^2/2}\,dz,$$

если b не равно иулю. Если b равно нулю, то

$$\frac{\log y_{ij}^{n} - na}{\sqrt{n}} \to 0$$

по вероятности при $n \to \infty$.

Не ясно, насколько типичен этот факт (асимптотическая нормальность), но очевидно, что он не имеет места для матриц общего характера. Можно взять, например, ортогональные матрицы, для которых известны соответствующие предельные теоремы (см. гл. 3). Как бы то ни было, было бы желательно привести этот результат в соответствие с общей теорией.

7.5.3. Чтобы проиллюстрировать понятие случайного спектра, укажем на два частных случая.

Первый относится к *многомерному статистическому анализу*. В этой теории большое внимание уделяется иекоторым случайным матрицам, составленным из оценок дисперсий, ковариаций и коэффициентов корреляции. Элементы будут иметь распределеиие Уишарта или иногда нормальное распределение. В статистическом анализе используются собственные значения $\lambda_1, \lambda_2, \ldots, \lambda_p$, они образуют случайные величины с совместным распределением вероятностей, плотность которого иногда может быть явно вычислена. Заметим, что мы должны иметь возможность как-то идентифицировать собственные значения, например упорядочивая их: $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_p$. Для того чтобы привести пример подобной функции плотности, предположим, что x_{ij} независимы, если не считать условия симметрии, так что x_{ij} и $x_{\alpha\beta}$ независимы, если $(i, j) \neq (\alpha, \beta)$ и $(i, j) \neq (\beta, \alpha)$, $x_{ij} = x_{ji}$, и что $x_{ij} = N(0, \sigma_{ij}^2)$, где $\sigma_{ij}^2 = 1/2$, если $i \neq j$, и $\sigma_{ii} = 1$. Тогда совместная функция плотности собственных значений определяется выражением

$$\frac{\exp\left(-\frac{1}{2}\sum_{1}^{p}\lambda_\nu^2\right)}{2^{p/2}\prod_{1}^{p}\Gamma\left(\frac{p+1-i}{2}\right)}\prod_{i<j}(\lambda_i-\lambda_j)$$

в области $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_p$ и равна нулю вне ее. Можно вычислить распределение простых функций собственных значений, таких, как определитель, равный $\lambda_1\lambda_2 \ldots \lambda_p$ (см. замечания 7.5.3).

С приведенным выше выражением для плотности нелегко иметь дело, если p велико. Интересно поэтому отметить следующий асимптотический результат (см. замечания 7.5.3). Пусть x_{ij} независимы, если не считать условия симметрии, но имеют распределения, относительно которых можно только утверждать, что 1) они симметричны относительно нуля, 2) они имеют дисперсию, равную единице, и 3) они имеют равномерно ограниченные моменты любого конечного порядка $|\mathbf{E}x_{ij}^r| \leqslant A_r < \infty$, $r = 1, 2, \ldots$. Рассмотрим тогда симметричные случайные матрицы

$$M_n = \frac{1}{\sqrt{n}}\{x_{ij};\ i,\ j = 1,\ 2,\ \ldots,\ n\}$$

с действительными собственными значениями $\lambda_1, \lambda_2, \ldots, \lambda_n$. Образуем случайные величины

$$\mu_p^{(n)} = \frac{1}{n}\sum_{\nu=1}^{n}\lambda_\nu^p = \frac{1}{n}\operatorname{tr} M_n^p = \frac{1}{n^{1+p/2}}\sum x_{i_1 i_2} x_{i_2 i_3} \ldots x_{i_p i_1},$$

где все индексы $i_1, i_2, \ldots, i_p$ пробегают значения $1, 2, \ldots, n$. Легко видеть, что если p — нечетное, то $\mathbf{E}\mu_p = 0$, в силу симметрии распределений. Если p — четное, скажем, $p = 2\nu$, то

$$\mathbf{E}\mu_{2\nu}^{(n)} = \frac{1}{n^{1+\nu}} \sum \mathbf{E} x_{i_1 i_2} x_{i_2 i_3} \ldots x_{i_{2\nu} i_1}.$$

Ясно, что многие члены в написанной выше сумме равны нулю в силу предположений о независимости и о том, что каждый множитель имеет среднее значение, равное нулю. Нетрудно доказать, что для больших значений n главный член в асимптотическом разложении суммы будет иметь вид $c_{2\nu} \cdot n^{\nu+1}$, а остаток имеет меньший порядок. Следовательно, $\lim\limits_{n \to \infty} \mathbf{E}\mu_{2\nu}^{(n)} = c_{2\nu}$; $\nu = 1, 2, \ldots$ Вычисление c является более хитрой комбинаторной задачей. Было показано (см. замечания 7.5.3), что $c_{2\nu} = (2\nu)!/\nu!\,(\nu+1)!$ Но характеристическая функция с такими моментами (и нечетными моментами, равными нулю) является целой функцией

$$\sum_{\nu=0}^{\infty} c_{2\nu} \frac{(iz)^{2\nu}}{(2\nu)!} = \sum_{\nu=0}^{\infty} \frac{(-z^2)^{\nu}}{\nu!\,(\nu+1)!} = \frac{1}{2} J_1(z) =$$

$$= \frac{2}{\pi} \int_{-1}^{1} \exp(2izu) \sqrt{1-u^2}\, du.$$

Но отсюда вытекает, что для любого полинома $c(x)$ имеем

$$\lim_{n\to\infty} \mathbf{E} \frac{1}{n} \sum_{\nu=1}^{n} c(\lambda_\nu) = \frac{1}{\pi} \int_{-2}^{2} c(u) \sqrt{1-\frac{u^2}{4}}\, du.$$

Теперь можно было бы попытаться использовать это для того, чтобы показать, что для любого интервала $(\alpha, \beta) \subset \subset (-2, 2)$ имеем

$$\lim_{n\to\infty} \mathbf{E} \left[\frac{1}{n} \cdot \text{число } \lambda_\nu \in (\alpha, \beta) \right] = \frac{1}{\pi} \int_{\alpha}^{\beta} \sqrt{1-\frac{u^2}{4}}\, du.$$

Функция мощности (см. разд. 7.2) должна тогда иметь предел

$$s(u) = \frac{1}{\pi} \sqrt{1-\frac{u^2}{4}}$$

и не зависеть от вида распределения x до тех пор, пока не будут выполнены сделанные предположения (см. замечания 7.5.3).

Вместо этого мы используем следующее рассуждение, которое не только описывает среднее поведение спектра, но и говорит нам кое-что относительно действительного асимптотического вида стохастического спектра. Для того чтобы доказать это более содержательное предложение, рассмотрим не только среднее поведение, но и дисперсии величин $\mu_{2\nu}^{(n)}$. Имеем

$$\mathbf{E}\,(\mu_{2\nu}^{(n)})^2 = \frac{1}{n^{2+2\nu}} \sum \mathbf{E} x_{i_1 i_2} x_{i_2 i_3} \ldots x_{i_{2\nu} i_1} x_{j_1 j_2} x_{j_2 j_3} \ldots x_{j_{2\nu} j_1}.$$

Чтобы вычислить эту сумму, зафиксируем индексы i_1, $i_2, \ldots, i_{2\nu}$ и посмотрим, что будет происходить, когда остальные индексы меняются. Но тогда эти члены распадаются на

$$\mathbf{E} x_{i_1 i_2} x_{i_2 i_3} \ldots x_{i_{2\nu} i_1} \cdot \mathbf{E} x_{j_1 j_2} x_{j_2 j_3} \ldots x_{j_{2\nu} j_1},$$

за исключением случая, когда хотя бы одна из пар (j_1, j_2), $(j_2, j_3), \ldots, (j_{2\nu}, j_1)$ равна по крайней мере одной из пар (i_1, i_2), $(i_2, i_3), \ldots, (i_{2\nu}, i_1)$, (i_2, i_1), $(i_3, i_2), \ldots, (i_{2\nu}, i_1)$. Но подобная ситуация встречается только в пренебрежимом числе случаев, так что действительно

$$\lim_{n\to\infty} \mathbf{E}\,(\mu_{2\nu}^{(n)})^2 = c_{2\nu}^2,$$

так что $\mathbf{D}\,(\mu_{2\nu}^{(n)}) \to 0$, и все моменты $\mu_{2\nu}^{(n)}$ сходятся в среднем к соответствующим пределам $c_{2\nu}$. Это имеет место также для моментов нечетного порядка. Рассмотрим теперь функции распределения $F^{(n_i)}(u)$ для произвольного фиксированного значения u и для произвольной подпоследовательности $n_1, n_2, n_3, \ldots$ натуральных чисел. Под $F^{(n)}(u)$ мы понимаем функцию распределения со скачками $1/n$ в точках $\lambda_1, \lambda_2, \ldots, \lambda_n$. Но все $\mu_{2\nu}^{(n)}$ сходятся по вероятности, так что из $\{n_i\}$ можно выбрать подпоследовательность $\{n_i'\}$, такую, что $\mu_2^{(n_i')}$ сходится почти наверное, затем подпоследовательность $\{n_i''\} \subset \{n_i'\}$, такую, что $\mu_4^{(n_i'')}$ сходится почти наверное и т. д. Диагональная процедура дает нам последовательность $n'_1, n''_2, \ldots$, для которой все моменты сходятся почти наверное. Но предельные моменты соответствуют, как мы видели, некоторой целой характеристической функции, откуда вытекает, что $F^{(k)}(u)$ сходится почти наверное к $F(u)$, когда k пробегает некоторую подпоследовательность последовательности $\{n_i\}$. Так как этого можно добиться, исходя из любой заданной последовательности $n_1, n_2, \ldots$, то $F^{(n)}(u)$; $n = 1, 2, \ldots$, сходится по вероятности к $F(u)$. Стохастический спектр стремится в этом случае к фиксированному спектру. Это может показаться удивительным, но, по-видимому, имеет место некоторое эргодическое свойство, связанное с независимостью распределений элементов матрицы. Если это так, то мы вправе ожидать, что сходимость того же вида имеет место и в более общей обстановке.

7.5.4. Рассмотрим дробно-линейное преобразование $1/(\alpha - x)$, $-1 < x < 1$, с точки зрения разд. 7.3. Для параметра α задается распределение вероятностей на множестве $|\alpha| > 1$. Так как

$$\left|\frac{1}{\alpha - x_1} - \frac{1}{\alpha - x_2}\right| \leqslant \frac{|x_1 - x_2|}{(|\alpha| - 1)^2},$$

то можно выбрать величину $\varrho(\alpha) = 1/(|\alpha| - 1)^2$ и потребовать, чтобы $\mathbf{E}\varrho(\alpha) < 1$. Это, конечно, справедливо, если $|\alpha| > 2$ с вероятностью единица.

Настоящая ситуация не соответствует в точности теореме 7.3.2 (здесь Z не является банаховым пространством), но можно показать с помощью рассуждения, аналогичного проведенному в разд. 7.3, что можно образовать сходящуюся непрерывную дробь

$$y_n = \cfrac{1}{\alpha_n - \cfrac{1}{\alpha_{n-1} - \cfrac{1}{\ldots}}},$$

где α независимы и имеют одинаковое распределение. Сходимость в указанном частном случае может быть установлена непосредственно с помощью теоремы Ворпицкого из общей теории непрерывных дробей (см. замечания 7.5.4).

Для y_n имеем рекуррентную формулу

$$y_{n+1}=\frac{1}{a_{n+1}-\frac{1}{y_n}},$$

которую можно использовать для того, чтобы выписать интегральное уравнение для функции распределения y.

Важным случаем, приводящим нас к случайным непрерывным дробям, является разностное уравнение второго порядка

$$z_{n+1}+a_n z_n+b_n z_{n-1}=0,$$

где a_n, b_n — случайные величины с распределениями, не зависящими от n и такими, что (a_n, b_n) не зависит от (a_m, b_m), если $m\neq n$. Разделив на z_n, получим, полагая $y_n = z_n/z_{n-1}$,

$$y_{n+1}=-a_n-\frac{b_n}{y_n},$$

что дает нам разложение в непрерывную дробь для y_{n+1}.

Более существен следующий пример. Пусть X состоит из действительных $k\times k$-матриц, рассматриваемых как линейные операторы в R^k с обычной нормой $\|x\|$. Предположим, что вероятностная мера определена на X так, что $\mathbf{E}\|x\|=c<1$. Введем случайные элементы

$$y_n=x_n+x_nx_{n-1}+x_nx_{n-1}x_{n-2}+\ldots=$$
$$=x_n\{e+x_{n-1}[e+x_{n-1}(e+\ldots)]+\ldots\}.$$

Сходимость этого выражения следует из рассуждений разд. 7.3. Имеем, очевидно,

$$y_{n+1}-x_{n+1}y_n=x_{n+1},$$

что представляет собой линейное разностное уравнение первого порядка со случайными коэффициентами x_{n+1} и случайными возмущениями в правой части. Случайные уравнения и процессы такого типа (принимающие значения из банаховой алгебры), могут оказаться очень полезными во многих приложениях. Это легко модифицировать для более общих случаев.

Проблема, которой много занимались (см. замечания 7.5.4), связана со случайными интегральными уравнениями, например фредгольмовского типа. Здесь случайность присутствует в ядре уравнения, в заданной функции для неоднородного уравнения или в области интегрирования. Аналогично дело обстоит и для других стохастических линейных операторов. Часто можно доказать существование случайных решений. Значительно более трудно, по-видимому, определить распределение вероятностей решения или даже найти его моменты. До сих пор это сделано только для очень частных случаев, и эта задача, без сомнения, привлечет большое внимание в будущей работе.

ОБЗОР

Теория вероятностей на алгебраических структурах несомненно имеет практическое значение уже в ее теперешней форме и может оказаться привлекательной для некоторых математиков. Многие известные результаты находятся в процессе кристаллизации в общую теорию. Но нельзя скрывать того, что в настоящее время в этой теории много пробелов и, к сожалению, некоторые из этих пробелов приводят к серьезным трудностям, когда дело доходит до приложений. Объясним это подробнее.

Но сначала рассмотрим несколько менее серьезные (по мнению автора) недостатки теории. Мы не пытались изложить теорию в наиболее общем виде или дать необходимые и достаточные условия и т. д. Мы не делали этого даже тогда, когда это можно было сделать с помощью уже установленных результатов и тем более тогда, когда это требовало серьезных усилий. Возьмем, например, сепарабельность рассматриваемых групп. Можно было бы, конечно, развивать анализ Фурье распределений вероятностей на компактной группе, не постулируя второй аксиомы счетности. Сепарабельность банаховых пространств X и X^* также может оказаться несущественной для результатов, подобных тем, которые были изучены. Можно было бы попытаться заменить измеримость по Борелю более общим понятием. Возможно, что некоторые утверждения, сделанные относительно стохастических групп, имеют место и без предположения локальной компактности, но здесь мы можем встретиться с препятствием, которое преодолеть труднее. Проблемы подобного рода, конечно, заслуживают внимания тех математиков, которым они по душе. Конечно, было бы приятно иметь подход к анализу Фурье распределений вероятностей на очень общих алгебраических и топологических структурах, включающих все те структуры, которые были изучены и, возможно, некоторые другие. Существует, однако, опасность того, что столь общий метод не даст слишком много нового и приведет только к очень бедной теории вероятностей.

Даже при сделанных предположениях существует много сложных вопросов, остающихся без ответа. Достаточно упомянуть о виде теоремы непрерывности на локально компактной группе или на банаховом пространстве. По сравнению с тем, что нам известно на эту тему в случае действительной прямой, положение неудовлетворительно. Аналогично дело обстоит с общими однородными процессами и бесконечно делимыми распределениями вероятностей: ощущается необходимость в детальной характеризации и полном описании. Нужен также метод описания преобразований Фурье распределения вероятностей в виде критерия, который можно было бы использовать на практике.

Даже если не удается получить полные ответы на такие вопросы, мы имеем по крайней мере частнчную информацию в этой области. Перейдем теперь к более трудным и насущным вопросам. На действительной прямой мы имеем богатую информацию относительно предельных распределений, иногда в неявной, иногда в законченной форме. На полугруппах, группах и т. д. положение много хуже. Возьмем группы матриц, которые среди стохастических групп имеют особенно большой практический интерес. Предположим, что в некоторой частной ситуации мы можем показать, что некоторое распределение может быть аппроксимироваио каким-то предельным распределением, скажем, нормальным. Как это можно использовать практически? Если только мы не имеем дела с особенно простыми ситуациями, у нас нет подхода для получения явного выражения для этого распределения. Может оказаться возможным вычисление нужной вероятности путем численного решения данного параболического уравнения. Если это оказывается возможным, то вряд ли мы можем пойти дальше и вычислить эти распределения для многих значений параметров, так как число их может быть огромным (см. гл. 7). Поэтому мы должны искать методы выражения этих распределений как параметрических семейств таких основных распределений, которые могут быть вычислены раз и навсегда. В настоящее время ни одного такого метода не известно. Можно было бы возразить, что подобного рода трудности и следовало ожидать. Наше знание нормальных распределений в конечномерных евклидовых пространствах является результатом большой работы, и нужно быть готовым к детальному изучению каждой из многих других стохастических групп, представляющих для нас интерес, для того чтобы прийти к численно полезным результатам.

В некоммутативной структуре мы можем столкнуться с той трудностью, что множество предельных распределений (для различно распределенных компонент) не замкнуто относительно композиции. Пусть, скажем, мы изучаем произведения вида

$$u_n = s_1^{(n)} s_2^{(n)} \dots s_n^{(n)} t_1^{(n)} \dots t_n^{(n)},$$

где все независимые случайные элементы $s_\nu^{(n)}$ имеют одно и то же распределение P_n и все элементы $t_\nu^{(n)}$ имеют распределение Q_n. Мы знаем, какие условия надо наложить на P_n и Q_n для того, чтобы

$$s_1^{(n)} s_2^{(n)} \dots s_n^{(n)} \to L_1, \quad n \to \infty,$$

$$t_1^{(n)} t_2^{(n)} \dots t_n^{(n)} \to L_2, \quad n \to \infty.$$

Рассмотрим случай, когда предельные распределения L_1 и L_2 принадлежат к сложному пуассоновскому типу. Известно, что $L_1 * L_2$ не обязано быть сложным пуассоновским распределением, так что предельное распределение произведения u_n может иметь другой вид. Это противоположно тому, что имеет место на действительной прямой, где предельные распределения часто одни и те же и для равных и для различных компонент. В нашей более общей обстановке случай различных компонент вряд ли вообще рассматривался, и здесь мы можем встретиться с серьезными трудностями.

Рассмотрим, наконец, следующую основную проблему для некомпактных стохастических групп. Производя выборку из распределения $P \in \mathcal{P}(G)$, образуем произведение $\gamma_n = g_1 g_2 \dots g_n$ и будем интересоваться аппроксимациями распределения P^{n*} произведения γ_n. Заметим, что распределение P отдельных множителей сохраняется фиксированным и не стремится к δ_e с увеличением объема выборки. Ни одна из имеющихся предельных теорем не может быть непосредственно применена в этом случае, и мы должны подумать о каком-то новом подходе. Хотелось бы получить результат вроде следующего. Существует последовательность $M_1, M_2, \ldots$ взаимно однозначных отображений группы G на себя, такая, что элементы $M_n\gamma_n$ сходятся по распределениям к некоторому невырожденному $Q \in \mathcal{P}(G)$. Приходим к мысли о выборе M_n как автоморфизмов и использовании затем анализа Фурье, но в настоящее время сомнительно, чтобы это привело к успеху. Это, возможно, самая важная из проблем, остающихся открытыми для стохастических групп.

Больше всего мы нуждаемся в дополнительном опыте обращения с конкретными стохастическими структурами. Мы можем надеяться достичь этого при применениях теории к специальным задачам.

ЗАМЕЧАНИЯ

1.2.1. Мы предполагаем известными основные понятия теории вероятностей (см., например, Лоэв [1]). Нам понадобятся также результаты общей теории меры, скажем, в объеме книги Халмоша [1].

1.2.2.$_1$. Известно, что функция $\varphi(z)$ тогда и только тогда является характеристической функцией устойчивого распределения, когда

$$\log\varphi(z) = i\gamma z - c\,|z|^{\alpha}\left[1 + i\beta\frac{z}{|z|}\,\omega(z,\alpha)\right],$$

где константы α, β, γ, c удовлетворяют неравенствам

$$0 < \alpha \leqslant 2, \quad -1 \leqslant \beta \leqslant 1, \quad -\infty < \gamma < +\infty, \quad c \geqslant 0,$$

а

$$\omega(z,\alpha) = \begin{cases} \operatorname{tg}\frac{\pi}{2}\alpha, & \text{если } \alpha \neq 1, \\ \frac{2}{\pi}\log|z|, & \text{если } \alpha = 1 \end{cases}$$

(см. Гнеденко и Колмогоров [1], стр. 177).

1.2.2.$_2$. Подходящим методом изучения распределения некоторых функционалов частичных сумм случайных величин является принцип инвариантности Эрдёша, Каца и др. (см. Донскер [1]). Пусть $S_\nu = x_1 + x_2 + \ldots + x_\nu$ есть частичная сумма независимых и одинаково распределенных случайных величин x_i со средними значениями 0 и дисперсией 1. Если F — достаточно хорошая функция от $S_1, S_2, \ldots, S_n$, то можно показать, что $F(S_1, S_2, \ldots$ $\ldots, S_n)$ имеет асимптотически то же распределение, что и F^* $(x, (t))$, где F^* — некоторый функционал в C (0, 1), тесно связанный с F, а $x(t)$ — винеровский процесс.

1.3.2. Эта проблема рассматривается в книге Гренандера и Розенблата [1], стр. 51.

1.3.3. Эта проблема рассматривается в работе Перрэна [1], стр. 1.

1.3.4. Неупорядоченные линейные структуры исследуются во многих работах по физике. Укажем только на особенно интересную работу Дайсона [1].

1.4.3. Это понятие математического ожидания является обобщением понятия, естественного в R^k. Пусть $x = (x_1, x_2, \ldots, x_k)$ — случайный k-мерный вектор. Обычно определяют математическое ожидание вектора x как вектор с компонентами $(\mathbf{E}x_1, \mathbf{E}x_2, \ldots, \mathbf{E}x_k)$, если они существуют. В бесконечномерном векторном пространстве, скажем, имеющем базис, существует дополнительная трудность, заключающаяся в том, что в пространстве может не оказаться вектора с компонентами $\mathbf{E}x_\nu$, $\nu = 1, 2, \ldots$

1.4.7. Изучались также распределения вероятностей в однородных пространствах. Вулл [1] рассматривал однородные случайные процессы со значениями в таком пространстве. Хотя это и относится к нашей теме, но выходит за рамки книги и здесь не будет обсуждаться.

2.1. Сепарабельным мы называем пространство, имеющее счетный базис (вторая аксиома счетности). Когда мы имеем дело с метрическими пространствами, это эквивалентно существованию счетного всюду плотного множества (см., например, Данфорд и Шварц [1], стр. 32); обычно это свойство и принимается за определение сепарабельности для таких пространств.

Сепарабельность не всегда существенна для последующих результатов, и при дальнейшем развитии теории она, наверное, может быть заменена значительно более слабым предположением. Условие сепарабельности удобно, так как оно позволяет избежать некоторых трудностей, связанных с вопросами измеримости. К тому же гильбертовы пространства $L_2(G)$, $\mathscr{H}$ и т. д. становятся сепарабельными; множество неприводимых унитарных представлений компактной группы становится счетным и т. д.

Можно использовать другие σ-алегбры в качестве области для вероятностной меры, а не ту, которая использована в тексте. Единственным требованием, налагаемым на эти алгебры, является их замкнутость относительно используемых алгебраических операций: если E и F принадлежат σ-алгебре, то их (алгебраическое) произведение EF также должно принадлежать ей.

2.1.$_1$. Хотя это не всегда оговаривается в тексте, часто будет иметься в виду вероятностное пространство $(\Omega, \mathscr{F}, \mu)$, и случайный элемент s представляется тогда в виде $s = s(\omega)$, где $s(\omega)$ — (борелевское) измеримое отображение: прообраз борелевского множества из S должен принадлежать $\mathscr{F}$.

В этой связи стоит отметить следующую проблему. Если мы хотим рассматривать бесконечную систему $\{s_t;\ t \in T\}$ случайных элементов, то можно ли построить вероятностную меру на прямом произведении S^T, исходя из конечномерных (частных) распределений? Другими словами, проходит ли колмогоровская конструкция (первоначально проведенная в R^T) в этом, более общем случае?

Такие вопросы только по временам обсуждаются в тексте. Заметим только, что эта конструкция проходит, если S — локально компактное хаусдорфово пространство (см. Пакшираджан [1]) или если S — полное метрическое пространство (см. Шпачек [2]).

Предположение регулярности, которое делается в тексте, не очень важно и, в действительности, не всегда используется в книге. Иногда, однако, оно удобно, например для того, чтобы получить теорему Рисса о представлении в ее обычной формулировке. Действительно, пока мы имеем дело с сепарабельными локально компактными пространствами, каждая борелевская мера есть мера Бэра и, следовательно, регулярна (см. Халмош [1], стр. 223).

Индекс ω в символе $L_\omega(S)$ указывает на бесконечную точку. Если дано локально компактное пространство S, то мы можем превратить его в компактное пространство S_ω, добавив новый элемент ω и определяя систему открытых множеств в S_ω как систему, состоящую из всех открытых множеств в S и из дополнений компактных подмножеств в S. Если (в случае, когда S — полугруппа) мы также положим по определению $s\omega = \omega s = \omega$, то S_ω будет полугруппой, но не обязательно топологической, даже если S — топологическая полугруппа. Нечего и говорить о том, что группа теряет групповые свойства при добавлении к ней ω. Это сильно снижает пользу описанной процедуры (одноточечного компактного расширения).

2.1.$_2$. Более подробное изложение можно найти в книге Халмоша [1].

2.1.$_3$. Действительно, рассмотрим P как линейные функционалы с нормой 1 в пространстве $L_\omega(S)$. Тогда L-слабая сходимость есть не что иное, как слабая * [1]) сходимость. При этом компактность $\mathcal{P}(S)$ следует из общих принципов (см., например, Хилле и Филлипс [1], стр. 50). Утверждение это может быть проверено и непосредственно по аналогии со случаем действительной прямой с использованием факта существования счетного базиса.

2.2.$_1$. См. Стромберг [1].

2.2.$_2$. Пусть имеют место слабые сходимости $P_n \to P$, $Q_n \to Q$. Тогда для любой непрерывной и ограниченной функции $f(s)$ имеем по теореме Фубини

$$\int_S f(s) P_n * Q_n(ds) = \int_S \int_S f(st) P_n(ds) Q_n(dt) = \int_S g_n(t) Q_n(dt),$$

где

$$g_n(t) = \int_S f(st)\, P_n(ds).$$

Для любого t имеем

$$g_n(t) \to g(t) = \int_S f(st) P(ds).$$

[1]) Слабая * топология в X^* в русском переводе книги Хилле и Филлипса [1] называется X-топологией. — *Прим. перев.*

Но последовательность $g_n(t)$ равностепенно непрерывна на любом компактном множестве C. Действительно, для любого $\varepsilon > 0$ можно выбрать компактное множество C_ε, такое, что $P_n(C_\varepsilon) \geqslant 1 - \varepsilon$. Следовательно,

$$|g_n(t_2) - g_n(t_1)| \leqslant 2\varepsilon \sup_s |f(s)| + \max_{s \in C_\varepsilon} |f(st_2) - f(st_1)|,$$

откуда вытекает равностепенная непрерывность. Отсюда и из $g_n(t) \to g(t)$ с помощью стандартных рассуждений получаем, что эта сходимость равномерна на C. Следовательно,

$$\int_S f(s) P_n * Q_n(ds) - \int_S f(s) P * Q(ds) = \int_C [g_n(t) - g(t)] Q_n(dt) + \\ + \int_{S-C} [g_n(t) - g(t)] Q_n(dt) + \int_S g(t) [Q_n(dt) - Q(dt)].$$

Все три интеграла в правой части могут быть сделаны как угодно малыми, так что имеет место слабая сходимость $P_n * Q_n \to P * Q$.

Мы не пользуемся систематически тем, что $\mathscr{P}(S)$ — топологическая полугруппа. По-видимому, используя это обстоятельство, можно получить полезные результаты.

(См. также работу Калианпура [1], в которой рассматриваются подобные вопросы.)

2.2.3. Интуитивное содержание теоремы 2.2.2 ясно: случайный элемент $u = st$ полугруппы пробегает *в точности* $\overline{s(P) \cdot s(Q)}$, если пренебречь событиями нулевой вероятности. В компактном случае соответствующее равенство $s(P * Q) = s(P) \cdot s(Q)$ было доказано Розенблатом [1], а для компактных групп это соотношение отмечалось различными авторами.

2.2.4. Представляется естественной следующая терминология. Рассмотрим непрерывный однородный процесс P_t, $t \geqslant 0$, выборочные функции которого почти наверное непрерывны. Такой процесс мы называем броуновским движением на полугруппе. Распределение вероятностей $P \in \mathscr{P}(S)$ называется нормальным распределением, если оно может быть вложено в броуновское движение P_t,

$P = P_1$. Поскольку мы не изучаем свойства выборочных функций, этот подход не будет здесь разбираться подробно. Предполагается, что выполнены следующие свойства: а) если P и Q—два коммутирующих нормальных распределения, то $P * Q$ есть нормальное распределение; б) если P — нормальное распределение на локально компактной группе G и N — нормальный делитель группы G, то индуцированное распределение вероятностей на факторгруппе $F = G/N$ является нормальным (см. конец разд. 3.1).

2.3.$_1$. Читатель может предпочесть следующее несколько более ясное интуитивное определение сходимости по вероятности. Говорят, что случайные элементы s_n с распределениями вероятностей P_n сходятся по вероятности к постоянному элементу s_0, если $P_n(N) \to 1$ для любой окрестности N элемента s_0. Для того чтобы убедиться, что это эквивалентно слабой сходимости P_n к δ_{s_0}, предположим, что имеет место $P_n(N) \to 1$. Если $f(s)$ — произвольная непрерывная ограниченная действительная функция на S, то

$$\int\limits_S f(s) P_n(ds) = \int\limits_N + \int\limits_{S-N} f(s) P_n(ds).$$

Если выбрать окрестность N элемента s_0 столь малой, что $|f(s) - f(s_0)| < \varepsilon$, то первое слагаемое в правой части будет мало отличаться от $f(s_0)$, а второе от нуля, так что

$$\int\limits_S f(s) P_n(ds) \to f(s_0) = \int\limits_S f(s) \delta_{s_0}(ds).$$

Чтобы убедиться в обратном, возьмем произвольную окрестность N элемента s_0 и выберем непрерывную функцию $f(s)$, такую, что

$$0 \leqslant f(s) \leqslant 1, \quad f(s_0) = 1, \quad f(s) = 0 \text{ при } s \overline{\in} N,$$

используя, например, лемму Урысона. Тогда

$$1 = f(s_0) \leftarrow \int\limits_S f(s) P_n(ds) = \int\limits_N f(s) P_n(ds) \leqslant P_n(N) \leqslant 1,$$

так что $P_n(N) \to 1$.

2.3.$_2$. Напомним читателю некоторые основные понятия теории полугрупп.

Непустое подмножество $L \subset S$ называется левым идеалом, если $SL \subset L$. Аналогично определяется правый идеал. Если $I \subset S$ есть одновременно левый и правый идеал, то он называется двусторонним идеалом.

Элемент $s \in S$ называется идемпотентом, если $s \cdot s = s^2 = s$. Компактная полугруппа имеет по крайней мере один идемпотент. Элемент $0 \in S$ называется нулем, если $0s = s0 = 0$ для любого $s \in S$.

Будем говорить, что идемпотент s подчинен другому идемпотенту t, если $st = ts = s$. Далее, идемпотент s называется примитивным, если не существует нулевого идемпотента, подчиненного s. Полугруппа S называется вполне простой, если каждый идемпотент примитивен, и для каждого $s \in S$ существуют такие идемпотенты u и v, что $us = s = sv$. Представление $T \times X \times Y$, используемое далее в тексте, может быть найдено у Уоллеса [1].

Компактная полугруппа имеет единственный минимальный двусторонний идеал. Он обозначается K и называется ядром полугруппы S. Ядро является вполне простой подполугруппой. Оно может быть представлено в виде

$$K = \bigcap_i SiS,$$

где i пробегает все идемпотенты полугруппы S.

Компактная полугруппа, в которой из $us = ut$ вытекает $s = t$ и из $su = tu$ вытекает $s = t$ (закон сокращения), является группой.

Мы будем использовать соотношение $SIS = S$, где I есть множество всех идемпотентов, и предполагать, что $SS = S$ (см. Кох и Уоллес [1]).

Полезные сведения можно почерпнуть в работе Нумакура [1], содержащей среди других вещей и полные доказательства приведенных выше утверждений.

Теоремы 2.3.1 и 2.3.2. *а*) и *б*) принадлежат Розенблату [1] и Хеблу и Розенблату [1]. Теорема 2.3.3 может быть найдена в работе Гликсберга [1]. Рассуждение в доказательстве теоремы 2.3.2. *а*) принадлежит Клоссу [1].

П. Мартин-Лёф предложил следующее простое доказательство теоремы 2.3.1. Множество $\mathcal{P}(S)$ компактно в слабой топологии. Пусть Q — произвольная предельная точка последовательности π_n. Тогда легко показать, что $Q * P = P * Q = Q$. Если Q' — некоторая другая предельная точка, то $Q' = Q' * P = Q' * P^{n*} = Q' * \pi_n = Q' * Q = Q$, что доказывает наше утверждение.

Если P имеет предел по Чезаро π, то можно показать, что сложный пуассоновский процесс $P_t = \exp^* tP$ обладает свойством $P_t \to \pi$ при $t \to \infty$.

2.3.$_3$. Это делается с использованием обычным образом свойства равностепенной непрерывности (см., например, Данфорд и Шварц [1], стр. 289).

2.3.$_4$. Здесь можно использовать вариант эргодической теоремы о среднем (см. Данфорд и Шварц [1], стр. 705). Пусть T — линейный оператор, отображающий банахово пространство X на себя, такой, что

а) последовательность $\frac{1}{n}(T + T^2 + \ldots + T^n)\,x$ ограничена,

б) для любого $x \in X$ последовательность $x_n = \frac{1}{n}(T + T^2 + \ldots + T^n)\,x$ имеет подпоследовательность, слабо сходящуюся к некоторому $\overline{x}$, и $\frac{1}{n} T^n x$ сходится к нулю.

Тогда x_n сходится к $\overline{x}$. Если T_0 обозначает отображение $x \to \overline{x}$, то $TT_0 = T_0T = T_0^2 = T_0$, $\| T_0 \| \leqslant C$. (Условие б) может быть ослаблено.)

Оператор T_0 — оператор проектирования на линейное многообразие $\{x \mid Tx = x\}$ неподвижных точек отображения T.

Теорема 2.3.1. принадлежит Розенблату [1], у которого читатель найдет также доказательство того, что $s(\pi)$ есть ядро полугруппы S.

2.3.$_5$. Для того чтобы убедиться в этом, заметим, что для любого s множество sK является идеалом, так как $sKS \subset sK$, $SsK = sSK \subset sK$, и оно содержится в K, так как $sK = \bigcap\limits_i sSiS \subset \bigcap\limits_i SiS = K$, так что $sK = K$. Отсюда вытекает существование единицы и единственного обратного элемента в K, так что K является группой.

Теоремы 2.3.3 и 2.3.4 принадлежат Гликсбергу [1] и Розенблату [1]. Укажем одну очень специальную предельную теорему. Если S — компактная топологическая полугруппа с нулем 0 ($0s = s0 = 0$ для всех $s \in S$) и если $0 \in s(P)$, то $P^{n*} \to \delta_0$. Доказательство предоставляется читателю.

2.3.$_6$. В этой книге мы все время будем встречаться с полугруппами линейных операторов T_t, $t > 0$, в которых $T_a T_b = T_{a+b}$ и которые в некотором смысле непрерывны. Если оператор T_t непрерывен в равномерной операторной топологии, т. е.

$$\lim_{h \downarrow 0} \| T_h - I \| = 0,$$

то инфинитезимальный оператор

$$A = \lim_{h \downarrow 0} \frac{T_h - I}{h}$$

существует как ограниченное линейное преобразование и $T_t = \exp(tA)$. Если T_t непрерывен в сильной топологии, т. е.

$$\lim_{h \downarrow 0} \| T_h x - x \| = 0 \text{ для любого } x \in X,$$

то существует предел $A_h = \frac{T_h - I}{h} x$ для элементов x, образующих плотное подмножество множества X. Предел A является, вообще говоря, неограниченным оператором. Тогда мы можем утверждать только, что

$$T_t x = \lim_{h \downarrow 0} \exp(tA_h) x.$$

Более подробные сведения по этому предмету читатель найдет в книге Хилле и Филлипса [1].

2.3.$_7$. Нелегко понять, что́ в действительности означает существование такого натурального m и насколько это предположение существенно для дальнейших результатов. Результаты, приведенные в тексте, принадлежат Хьюитту и Цуккерману [1], в работе которых читатель может найти более подробное обсуждение предельных результатов и бесконечных произведений элементов стохастической полугруппы.

3.1. Основы теории топологических групп можно найти в книгах Понтрягина [1] и Вейля [1].

3.1.$_1$. Мера, инвариантная справа, связана с мерой, инвариантной слева, соотношением

$$\nu(dg) = \frac{1}{\Delta(g)} \mu(dg)$$

с точностью до произвольного нормирующего множителя. Действительно, для любого компактного множества $E \subset G$ имеем

$$\int\limits_{g \in Eh} \frac{1}{\Delta(g)} \mu(dg) = \int\limits_{u \in E} \frac{1}{\Delta(u)\Delta(h)} \mu(duh) = \int\limits_{u \in E} \frac{1}{\Delta(u)} \mu(du),$$

так что функция множества

$$\int\limits_{E} \frac{1}{\Delta(g)} \mu(dg)$$

инвариантна справа, строго положительна и, следовательно, пропорциональна $\nu(E)$.

Так как $\Delta(g)$ непрерывна и отлична от нуля, то $\mu(E) = 0$ эквивалентно $\nu(E) = 0$ для любого компактного множества. Отсюда вытекает, что вероятностная мера, абсолютно непрерывная относительно одной из мер μ или ν, абсолютно непрерывна и относительно другой. Поэтому можно говорить об абсолютно непрерывных мерах (относительно мер, инвариантных справа или слева), опуская слова «справа» и «слева».

3.1.$_2$. Алгебра $L_1(G)$ изучается, например, в работе Хьюита [1].

3.1.$_3$. *Вероятностный оператор* T и некоторые его спектральные свойства были изучены Гренандером [2].

Вероятностные операторы

$$T_t f(g) = \int\limits_{G} f(gh) P_t(dh),$$

соответствующие непрерывному однородному процессу на компактной группе G, обладают некоторыми простыми свойствами:

а) T_t есть сильно непрерывная полугруппа в $C(G)$,

б) $T_t 1 = 1$,

в) T_t положительны, т. е. $T_t f(g) \geqslant 0$, если $f(g) \geqslant 0$ для всех $g \in G$,

г) T_t коммутируют с левыми переносами. Это легко проверить. Заметим также, что верно и обратное: оператор, обладающий этими свойствами, представим в указанном выше виде. Действительно, если выполнены условия а) и б), то для фиксированного t функционал $T_t f(e)$ может быть представлен в виде

$$T_t f(e) = \int_G f(h) P_t(dh), \quad P_t \in \mathcal{P}(G).$$

Но из свойства г) вытекает, что для любого левого переноса $Lx \to gx$ мы имеем

$$T_t f(g) = T_t L f(e) = L \int_G f(h) P_t(dh) = \int_G f(gh) P_t(dh).$$

Наконец, условие а) показывает, что P_t образует непрерывный однородный процесс на G.

3.1.$_4$. Форма эргодической теоремы, удобная для настоящей цели, приведена в кннге Данфорда и Шварца [1], стр. 717.

3.1.$_5$. Некоторые авторы используют термин «*непрерывный* гомоморфизм».

3.2. Теорема 3.2.1 принадлежит Уэнделу [1]. Можно сделать более подробные утверждения относительно примитивных идемпотентов в $\mathcal{P}(G)$ (см. замечания 2.3.$_2$) опять же в предположении, что G компактна. Коллинз [1] показал, что P является примитивным идемпотентом тогда и только тогда, когда $s(P) = H$ есть максимальная собственная замкнутая подгруппа группы G и P — нормированная мера Хаара на H.

3.2.$_1$. Преобразование Фурье распределений вероятностей на компактных группах было впервые использовано в классической работе Кавада и Ито К. [1], содержащей также, по существу, теорему 3.2.4.

Лемма 3.2.1 может быть высказана в более приятной форме, а именно:

Для того чтобы некоторое $\hat{P}_r$, $r \neq 0$, *имело норму* 1, *необходимо и достаточно, чтобы носитель* $s(P)$ *содержался в замкнутой собственной подгруппе* A *группы* G *или в классе смежности по* A.

Доказательство. Предположим, что $s(P) \subset \subset g_0 A$. Тогда симметричное распределение вероятностей $Q = \bar{P} * P$ имеет носитель, содержащийся в A. Его преобразование Фурье $\hat{P}^* \cdot \hat{P}$ является эрмитовым, и его норма $\| \hat{P}^* \cdot \hat{P} \|$ равна наибольшему по абсолютной величине собственному значению λ матрицы $\hat{Q} = \hat{P}^* \cdot \hat{P}$. Но модуль λ не может быть меньше единицы для всех представлений $r \neq 0$, так как тогда последовательность Q^{n*} сходилась бы к нормированной мере Хаара μ на G (см. лемму 3.2.2), имеющей носителем G. Следовательно, существует z такое, что $\| z \| = 1$ и $\hat{P}^* \cdot \hat{P} z = \pm z$ (для некоторого представления), причем $\| \hat{P} \|$ должна равняться единице.

С другой стороны, предположим, что $\| \hat{P} \| = 1$. Тогда существует такое z, что $\| z \| = 1$ и $\| \hat{P} z \| = 1$, т. е.

$$\left\| \int\limits_G U(g)\, zP(dg) \right\| = 1.$$

Но отсюда вытекает, так же как и в основном тексте, что $s(P)$ содержится в некотором $g_0 A$.

Занумеруем все функции $u_{ij}^{(r)}(g)$, входящие в неприводимые унитарные представления: $u_0(g) \equiv 1$, $u_1(g)$, $u_2(g)$, ... Пусть задано множество матриц $A^{(r)}$ с элементами $a_0, a_1, \ldots,$ занумерованными, как выше. Является ли оно преобразованием Фурье некоторого $P \in \mathcal{P}(G)$? Вот ответ на этот вопрос.

Теорема. *Матрицы* $A^{(r)}$ *образуют преобразование Фурье некоторого распределения вероятностей на группе тогда и только тогда, когда* $a_0 = 1$ *и для всех значений* n *и* c_ν *имеет место неравенство*

$$\left| \sum_{\nu=0}^{n} c_\nu a_\nu \right| \leqslant \max_{g \in G} \left| \sum_{\nu=0}^{n} c_\nu u_\nu(g) \right|.$$

Доказательство. Это практически эквивалентно классическому результату Ф. Рисса. Необходимость условия очевидна. Достаточность легко доказывается при следующем определении линейного функционала L: если

$$f(g)=\sum_{\nu=0}^{n} c_\nu u_\nu(g),$$

определим

$$Lf=\sum_{\nu=0}^{n} c_\nu a_\nu.$$

Так как $|Lf| \leqslant \|f\|$ и так как $u_\nu(g)$ образуют равномерно полную систему в $C(G)$, то L может быть распространено на $C(G)$ с нормой, не превосходящей единицы, и, следовательно, должна существовать вполне непрерывная функция множества $P(E)$ с вариацией, не превосходящей единицы, такая, что

$$Lf=\int_G f(g)P(dg),$$

откуда вытекает, что

$$Lu_\nu=a_\nu, \quad A=\hat{P}.$$

Так как $L1=1$, эта функция множества является вероятностной мерой.

3.2.$_2$. Сошлемся на Стромберга [1] по поводу модификации утверждения Кавада — Ито и на связанные с этим исследования Клосса [1], [2], Урбаника [1], Главки [1] и Коллинза [2].

Если P абсолютно непрерывна и обладает плотностью, интегрируемой с квадратом, то можно использовать теорию ядер Гильберта — Шмидта. В этом случае Ривкиндом [1] было показано, что P^{n*} сходится к нормированной мере Хаара, если G связна.

3.2.$_3$. Это неравенство установлено Клоссом [2].

3.3. Теорема 3.3.2 была доказана Герглотцем для случая, когда $G=N$ есть аддитивная группа целых неотрицательных чисел, и Бохнером для случая, когда $G = R^1$. Общий случай рассматривался Райковым и Вейлем (см., например, Вейль [1], стр. 137).

Относительно изучения $\sup_\gamma |\hat{P}(\gamma)|$ см. работу Хьюита и Рубина [1].

3.3.$_1$. Теорема 3.3.3 и ее доказательство, приведенное в тексте, принадлежат Рудину [2].

3.3.$_2$. Можно использовать интеграл Хеллингера для расширения определения $C(P)$ и утверждения теоремы.

3.3.$_3$. Если величина $d = \sqrt{\int \|g\|^2 \mu(dg)}$, где μ — нормированная мера Хаара, конечна, то существует граница для всех $d(P)$, $d(P) \leqslant d$. Действительно, по теореме Фубини

$$d^2 = \int_G d^2 P(dg) = \int_G \int_G \varphi(h) \mu(dh),$$

где

$$\varphi(h) = \int_G \|g - h\|^2 P(dg),$$

так что

$$d^2 \geqslant \inf_h \varphi(h) = d^2(P).$$

3.3.$_4$. (Добавлено при корректуре.) Некоторые пробелы в тексте, связанные с безгранично делимыми распределениями и предельными теоремами были заполнены недавно. Интересующиеся читатели могут посмотреть работу Партасарати, Ранга Рао и Варадхана [2] для случая локально компактной коммутативной группы и работу Клосса [3] для случая компактной коммутативной группы (см. также Карналь [1]).

3.4. Конечные стохастические группы изучались Воробьевым [1] и Дворецким и Вольфовицем [1], стохастическая группа вращений окружности — Леви [1]. Бёге [1] представил безгранично делимые распределения на конечной группе. Он показал, что 1) экспоненциальное представление не обязательно единственно и что 2) класс безгранично делимых распределений замкнут относительно операции композиции тогда и только тогда, когда (конечная) группа коммутативна.

4.1. Весьма полезной справочной книгой по основам теории групп Ли является книга Кона [1] (см. также последнюю главу книги Понтрягина [1]).

4.1.$_1$. Некоторые авторы изучали процессы диффузионного типа на дифференцируемых многообразиях (см., например, К. Ито [1] и С. Ито [1]).

4.2. Основная теорема этого раздела принадлежит Ханту [1]. Она обобщает классическое представление Леви — Хинчина безгранично делимых распределений на действительной прямой.

Рассуждения в разд. 4.2—4.4 тесно примыкают к выводам Ханта и Вена.

4.2.$_1$. Относительно доказательства см., напрнмер, Хант [1], стр. 269.

4.2.$_2$. Для того чтобы показать, что T_t отображает C'_k в C'_k, рассмотрим случай $k=1$. Тогда $Y'Tf = TY'f$, так как T коммутирует с левыми переносами. Следовательно, T отображает C'_k в C'_k и т. д. Также имеет место неравенство $\| Tf \|'_k \leqslant \| f \|'_k$. (См. Вен [2].)

4.2.$_3$. (См. Вен [1], [2].) Далее используется следующий результат, принадлежащий Вену. Рассмотрим последовательность операторов M_n описанного в тексте типа и характеризуемых постоянными $a_i^{(n)}$, $a_{ij}^{(n)}$ и мерами η_n, $n = 1, 2, \ldots$. Предположим, что

1) $\lim\limits_{n\to\infty} a_i^{(n)} = a_i$,

2) $\lim\limits_{\varepsilon\downarrow 0} \lim\limits_{n\to\infty} \left\{ a_{ij}^{(n)} + \frac{1}{2} \int\limits_{0<\varphi(g)\leqslant\varepsilon} x_i(g) x_j(g) \frac{\eta_n(dg)}{\varphi(g)} \right\}$ существует и равен a_{ij},

3) $\lim\limits_{n\to\infty} \eta_n = \eta$ всюду на $G_C - e$ и $\lim\limits_{n\to\infty} \eta_n(\omega) = \eta(\omega)$.

Тогда

$$\lim_{n\to\infty} \| (M_n - M) f \| = 0$$

для любой $f \in C_2$. Доказательство проводится следующим образом:

$$(M_n - M) f(g) = \sum_i (a_i^{(n)} - a_i) X_i f(g) +$$
$$+ \sum_{i,j} (a_{ij}^{(n)} - a_{ij}) X_i X_j f(g) + \int_{G_C - e} \Big[f(gh) - f(g) -$$
$$- \sum_i X_i f(g) x_i(h) \Big] \cdot \frac{\eta_n(dh) - \eta(dh)}{\varphi(h)} .$$

Разложим интеграл на две части, на одной из которых $0 < \varphi(h) \leqslant \varepsilon$, а на другой $\varphi(h) > \varepsilon$.

Норма второго интеграла стремится к нулю в силу условия 3). Первый интеграл в свою очередь может быть разложен

$$\int_{0<\varphi(h)\leqslant\varepsilon} \Big[f(gh) - f(g) - \sum_i X_i f(g) x_i(h) -$$
$$- \frac{1}{2} \sum_{i,j} X_i X_j f(g) x_i(h) x_j(h) \Big] \cdot \frac{\eta_n(dh) - \eta(dh)}{\varphi(h)} +$$
$$+ \frac{1}{2} \sum_{i,j} X_i X_j f(g) \int_{0<\varphi(h)\leqslant\varepsilon} x_i(h) x_j(h) \frac{\eta_n(dh) - \eta(dh)}{\varphi(h)} ,$$

где норма первого интеграла стремится к нулю, так как подинтегральная функция непрерывна и равна нулю на единичном элементе e. Второй интеграл уравновешивает в пределе выражение

$$\frac{1}{2} \sum_{i,j} (a_{ij}^{(n)} - a_{ij}) X_i X_j f(g)$$

(см. условие 2). Наконец, норма членов первого порядка в исходном выражении

$$\sum_i (a_i^{(n)} - a_i) X_i f(g)$$

стремится к нулю в силу условия 1).

4.2.₄. В настоящее время очень мало известно относительно нормальных распределений на группе Ли. Мы даже не знаем, является ли, вообще говоря, композиция двух нормальных распределений (соответствующих инфинитезимальным порождающим операторам M_1 и M_2) снова нормальной. Если M_1 и M_2 коммутируют, то это так.

Если P — нормальное распределение на группе Ли G, а N — замкнутый нормальный делитель, то распределение вероятностей P', индуцированное на факторгруппе $G' = G/N$ (см. конец разд. 3.1), также нормально. Чтобы убедиться в этом, вкладывают P в P_t, $P_1 = P$. Тогда мы можем определить P'_t соотношением

$$\int_{G'} f(g') P'_t(dg') = \int_G f[h(g)] P_t(dg)$$

для любой $f \in L(G')$, где h — естественный гомоморфизм $G \to G'$. Если $f \in L(G') \cap C_2(G')$, то инфинитезимальный порождающий оператор M' полугруппы P'_t задается соотношением

$$M'f(g') = Mf[h(g)],$$

где M — инфинитезимальный порождающий оператор P_t. Вспомним теперь выражение оператора M в терминах алгебры Ли Λ, натянутой на $X_1, X_2, \ldots, X_d$. Функция $f[h(g)]$ постоянна на классах смежности по подгруппе N, и инфинитезимальные преобразования, принадлежащие алгебре Ли n, связанной с N, будучи применены к этой функции, дают нуль. Другими словами, инфинитезимальный порождающий оператор M' может быть выражен членами первых двух порядков в инфинитезимальных преобразованиях, принадлежащих факторалгебре Λ/n. Но она изоморфна алгебре Ли группы $G/N = G'$ (см. Кон [1], стр. 132), и это заканчивает доказательство. Сравните это с тем, что говорится в разд. 4.5.1 относительно вида нормальной плотности, выраженной с помощью накрывающей группы.

4.2.₅. (См., например, Понтрягин [1], стр. 330.) Это простое следствие свойств унитарных представлений.

4.3.—4.4. Некоторые выводы в этих разделах только намечены, для того чтобы не затемнять математической сущности рассуждения. В частности, некоторые рассуждения, связанные с аппроксимацией, если и упомянуты в тексте, то только вскользь. Читатель, интересующийся полными доказательствами, должен обратиться к работам Вена [1], [2].

4.5.1. Конструкция процесса на накрывающей группе G^* принадлежит С. Ито. Понятие накрывающей группы изучается в книге Кона [1]. Относительно броуновских движений в группах Ли и их инфинитезимальных операторов см. К. Ито [2].

В интересной работе Мак-Кина [1] строится броуновское движение, исходя из броуновского движения в алгебре Ли. Его конструкция проводится для группы трехмерных вращений, но применима и для более общих групп Ли. (См. также разд. 7.1.)

Следующий пример принадлежит П. Мартин-Лёфу. Пусть $f(x, \sigma)$ — нормальная плотность на T^1:

$$f(x, \sigma)=\frac{1}{\sqrt{2\pi}\sigma}\sum_{\nu=-\infty}^{\infty}\exp\left[-\frac{(x-2\nu\pi)^2}{2\sigma^2}\right].$$

Определим

$$f_1(x)=f\left(x;\ \frac{\sigma}{\sqrt{2}}\right)+u\cos x,$$

$$f_2(x)=f\left(x;\ \frac{\sigma}{\sqrt{2}}\right)-2\,\frac{u\alpha}{u+2\alpha}\cos x,$$

где

$$\alpha=\int_{T^1} f\left(x;\ \frac{\sigma}{\sqrt{2}}\right)\cos x\,dx.$$

Если u достаточно мало, то $f_1(x)$ и $f_2(x)$ — плотности. Но

$$f_1*f_2(x)=f\left(x;\ \frac{\sigma}{\sqrt{2}}\right)*f\left(x;\ \frac{\sigma}{\sqrt{2}}\right)+$$
$$+u\left(1-\frac{2\alpha}{u+2\alpha}\right)f\left(x;\ \frac{\sigma}{\sqrt{2}}\right)*\cos x-$$
$$-2\,\frac{u^2\alpha}{u+2\alpha}\cos x*\cos x=f(x;\ \sigma),$$

так как

$$\cos x * \cos x = \frac{1}{2} \cos x,$$

$$f\left(x; \frac{\sigma}{\sqrt{2}}\right) * \cos x = \alpha \cos x.$$

Это показывает, что нормальное распределение на T^1 может быть представлено в виде композиции распределений, отличающихся от нормального.

4.5.2. См. Гренандер [3].

4.5.3. Более детальное изложение можно найти в работе Карпелевича, Тутубалина и Шура [1]. Эта же стохастическая группа рассматривалась Гетуром [1], в работе которого безгранично делимые распределения изучались с помощью преобразования Лежандра.

5.1. Исчерпывающим и прекрасно написанным введением в область локально компактных групп, унитарных представлений, положительно определенных функций и т. п. является книга Наймарка [1], а заинтересованный читатель может найти дополнительную информацию по этим вопросам в работах Годмана [1], [2]. Представление в виде прямого интеграла дано в работе Маутнера [1] и более полно в его же работе [2]. (См. также Хьюит [1], Люмис [1].)

5.1.$_1$. Реальное значение свойств P и P' еще не вполне понятно. Компактная группа обладает ими тривиальным образом. Используя лемму из книги Наймарка [1], стр. 375, можно показать, что локально компактная коммутативная группа также обладает этими свойствами. Для общих локально компактных групп ответ, однако, не известен. В литературе обсуждался вопрос, может ли функция, тождественно равная единице, быть равномерно аппроксимирована функциями вида $c(g) * \tilde{c}(g)$. Это связано с вопросом о норме вероятностного оператора T (см. разд. 3.1). Йосидзава [1] построил следующий пример, показывающий, что эта аппроксимация не всегда возможна при $c \in L_2(G)$. Пусть G — свободная группа с двумя образующими a и b. Тогда для всех достаточно малых $\varepsilon > 0$ не существует функции $c \in L_2(G)$, такой, что

$$\begin{cases} c * \widetilde{c}(e) = 1, \\ |c * \widetilde{c}(a) - 1| < \dfrac{\varepsilon^2}{2}, \\ |c * \widetilde{c}(b) - 1| < \dfrac{\varepsilon^2}{2}. \end{cases}$$

Действительно, если бы такая функция существовала; то при h, равном a или b, имело бы место неравенство

$$\int_G |c(g) - c(hg)|^2 \mu(dg) < \varepsilon^2,$$

так как при этих h

$$\int_G |c(g) - c(hg)|^2 \mu(dg) = 2 \int_G |c(g)|^2 \mu(dg) - 2 \operatorname{Re} c * \widetilde{c}(h) < \varepsilon^2.$$

Рассмотрим подмножество S_a группы G, состоящее из приведенных слов, начинающихся с некоторой (ненулевой) степени a, и аналогично определим S_b. Введем абсолютно непрерывную вероятностную меру

$$m(E) = \int_E |c(g)|^2 \mu(dg).$$

Тогда имеем

$$|m(aE) - m(E)| = \left| \int_E \{|c(ag)|^2 - |c(g)|^2\} \mu(dg) \right| \leqslant 2\varepsilon,$$

так что

$$m(aE) > m(E) - 2\varepsilon.$$

Возьмем теперь в качестве E множества S_a^*, aS_a^*, $a^2S_a^*$. Тогда $m(S_a^*) + m(aS_a^*) + m(a^2S_a^*) > 3m(S_a^*) - 4\varepsilon$. Но $m(S_a^*) + m(aS_a^*) + m(a^2S_a^*) \leqslant m(G) = 1$, так что

$$m(S_a^*) < \frac{1}{3} + \frac{4}{3}\varepsilon,$$

и аналогичным образом

$$m(S_b^*) < \frac{1}{3} + \frac{4}{3}\varepsilon.$$

Отсюда в силу того, что $S_a \bigcap S_b = \varnothing$ и $S_a^* \bigcup S_b^* = G$, получаем

$$1 = m(G) \leqslant m(S_a^*) + m(S_b^*) < \frac{2}{3} + \frac{8}{3}\varepsilon,$$

что не может иметь места для достаточно малых ε. Это доказывает невозможность равномерной аппроксимации функции, тождественно равной единице, функциями вида $c * \widetilde{c}(g)$, где $c \in L_2(G)$.

Такенути [1] изучил P^1-группы и охарактеризовал некоторые из них. (См. также Дьедонне [1], [2], [3].)

5.1.$_2$. См. Понтрягин [1], стр. 106.

5.1.$_3$. Утверждение в) принадлежит Годману [2].

5.2. В первоначальной форме вторая часть утверждения е) была высказана Гренандером [5], который предполагал, что группа должна обладать P-свойством, для того чтобы имела место теорема непрерывности. Лойнес [1] показал, что это предположение излишне. Первая часть утверждения е) принадлежит Годману [1].

Теорема 5.2.1 принадлежит Лойнесу [1].

5.2.$_1$. Локально компактная группа полна, т. е. в ней любая последовательность Коши имеет предел (см. Бурбаки [1], стр. 227).

Более подробное обсуждение различных видов стохастической сходимости в равномерных пространствах см. в работе Досса [2].

5.2.$_2$. Пакшираджан [1] изучал критерии сходимости для локально компактных коммутативных метрических групп. Среди других результатов он получил следующий: для того чтобы ряд $\sum_1^\infty g_n$ сходился почти наверное, необходимо и достаточно, чтобы произведение $\prod_1^\infty \mathbf{E}\gamma(g_n)$ сходилось для любого характера $\gamma \in G^*$.

5.2.$_3$. Два указанных типа сходимости эквивалентны также сходимости по вероятности (см., например, Лоэв [1], стр. 265).

5.2.$_4$. Это — стандартное аппроксимационное рассуждение (см. Наймарк [1], стр. 335 и 363).

5.3.$_1$. Относительно спектрального представления линейного оператора см., например, Рисс и Секефальви-Надь [1], стр. 304. Этот и следующие интегралы должны пониматься как берущиеся по полуоткрытому интервалу $[-\pi, \pi)$.

5.3.$_2$. О понятии и свойствах интеграла Бохнера см., например, Хилле и Филлипс [1], стр. 92.

5.3.$_3$. Такие операторнозначные функции, как $N(\lambda)$ и $M(\lambda)$, называются обобщенными разложениями единицы (см. Ахиезер и Глазман [1], стр. 359). При работе с такими функциями может оказаться полезным знание того, что они могут быть представлены следующим образом. Существуют такое гильбертово пространство $\mathscr{H}^+$, в которое вложено $\mathscr{H}$, и такое обычное разложение единицы $E^+(\lambda)$ в $\mathscr{H}^+$, что для любого $z \in \mathscr{H}$ имеем $N(\lambda)\, z = PE^+(\lambda)\, z$, где P — оператор проектирования $\mathscr{H}^+$ на $\mathscr{H}$.

5.4. См. Гренандер [5].

5.4.$_1$. Стоуновское представление группы унитарных операторов рассматривается, например, в книге Рисса и Секефальви-Надя [1], стр. 411.

5.5.2. Группа линейных преобразований $x \to \alpha x + \beta$ и ее неприводимые представления рассматриваются в книге Наймарка [1], стр. 342. Относительно других классических групп см. Гельфанд и Наймарк [1].

5.5.3. Важное соотношение $r = \sqrt{(2h-1)/h^2}$ было доказано Кестеном. В его работе [1] можно найти ряд относящихся сюда результатов. Укажем на следующее простое выражение для спектрального радиуса оператора M. Запишем M в спектральном представлении

$$M = \int\limits_{\lambda_1}^{\lambda_2} \lambda \, dE(\lambda),$$

где λ_1 выбрано столь большим, а λ_2 столь малым, насколько возможно, т. е. λ_1 и λ_2 — соответственно левый и правый экстремумы спектра. Конечно, спектр вполне определяется диагональными элементами $E_{ii}(\lambda)$ (мы представляем операторы $E(\lambda)$ в пространстве l_2). Вероятность вернуться из i в i за n шагов равна

$$M_{ii}^n = \int\limits_{\lambda_1}^{\lambda_2} \lambda^n \, d\, E_{ii}(\lambda).$$

Но все M_{ii}^n, а следовательно, все E_{ii} не зависят от i в силу своего вероятностного значения. Таким образом, спектральный радиус r определяется соотношением

$$r = \max(|\lambda_1|, |\lambda_2|) = \varlimsup_{n\to\infty} M_{ii}^n =$$

$$= [\text{радиус сходимости ряда } \sum_n M_{ii}^n x^n]^{-1}.$$

Кестен [2] показал также следующее. Пусть P — симметричная вероятностная мера на счетной группе $s(P)$, такая, что носитель $s(P)$ натянут на G. Тогда $\| T \| = 1$ тогда и только тогда, когда существует *полное банахово среднее значение* на G. Полное банахово среднее значение есть функционал L, определенный на пространстве $B(G)$ действительных ограниченных функций на G, такой, что

$$\begin{cases} L(c_1 f_1 + c_2 f_2) = c_1 L(f_1) + c_2 L(f_2), \\ \inf\limits_{g\in G} f(g) \leqslant L(f) \leqslant \sup\limits_{g\in G} f(g), \\ Lf(agb) = Lf(g), \end{cases}$$

где c_1 и c_2 — произвольные действительные постоянные; $f, f_1, f_2 \in B(G)$, а a и b принадлежат G.

6.1. В этой главе мы будем предполагать известными некоторые сведения относительно основ теории линейных пространств (см., например, первые две главы книги Хилле и Филлипса [1]).

Начало исследованию стохастических банаховых пространств положила работа Мурье [1], содержащая многие результаты этой главы. Очень хорошим источником информации является также работа Прохорова [1], особенно относительно сходимости распределений вероятностей в банаховом пространстве, и два глубоких исследования Прохорова [2] и Ле Кама [1]. (См. также Варадараджан [1].)

6.1.$_1$. Заметим прежде всего, что для того чтобы показать, что $\mathscr{L}$ содержит все открытые подмножества пространства X, достаточно доказать, что все шары в X принадлежат $\mathscr{L}$. Рассмотрим шар $S = \{x \mid \|x\| \leqslant 1\}$. Введем множество F всех $x^* \in X^*$, таких, что $\|x^*\| = 1$. Тогда

$$S = \bigcap_{x^* \in F} A_{x^*}, \quad A_{x^*} = \{x \mid |x^*(x)| \leqslant 1\}.$$

Действительно, $S \supset \bigcap A_{x^*}$ очевидно, а обратное включение доказывается следующим образом: если $x \in \bigcap A_{x^*}$ и $x^* \in F$, $x^*(x) = \|x\|$, то $\|x\| \leqslant 1$ и $x \in S$.

Выберем теперь последовательность $x_1^*, x_2^*, \ldots$, плотную в F в смысле слабой сходимости (это можно сделать в силу предположения сепарабельности), и определим

$$E_i = \{x \mid |x_i^*(x)| \leqslant 1\},$$
$$E = \bigcap_i E_i.$$

Очевидно, что $E \supset S$. С другой стороны, если x принадлежит E, но не принадлежит $S = \bigcap A_{x^*}$, то

$$\begin{cases} |x_i^*(x)| \leqslant 1 & \text{для всех } i, \\ |x^*(x)| > 1 & \text{для некоторого } x^* \in F. \end{cases}$$

Но это приводит к противоречию, так как это x^* может быть аппроксимировано с помощью x_i^*, и наше утверждение доказано.

В работе Крамера и Вольда [1] было доказано, что распределение вероятностей в R^k однозначно определено, если известна вероятностная масса, распределенная на каждом полупространстве $c_1x_1 + c_2x_2 + \ldots + c_kx_k < c$. Эти полупространства образуют семейство единственности.

До тех пор пока мы имеем дело с сепарабельными банаховыми пространствами, мы видим, что с полупространствами так же удобно работать, как и с открытыми множествами: σ-алгебры, порожденные этими двумя семействами, совпадают.

В нашем случае вероятностное пространство не является локально компактным, так что композиции не будут определяться точно так, как в гл. 2. Однако если X сепарабельно и x_1 и x_2 — независимые случайные элементы, то их сумма $x_1 + x_2$ есть также случайный элемент. Действительно, пусть $x_1(\omega)$ и $x_2(\omega)$ — независимые случайные элементы, определенные на вероятностном пространстве $(\Omega, \mathcal{A}, P)$. Как ясно из предыдущего, достаточно показать, что множества вида

$$\{\omega \mid x^* [x_1(\omega) + x_2(\omega)] \leqslant 1\} = \{\omega \mid x^* [x_1(\omega)] + x^* [x_2(\omega)] \leqslant 1\}$$

принадлежат σ-алгебре $\mathcal{A}$. Но $x^* [x_1(\omega)]$ и $x^* [x_2(\omega)]$ — обычные случайные величины, принимающие действительные значения, так что это верно для каждого $x^* \in X^*$. Таким образом, случайный элемент $x_1(\omega) + x_2(\omega)$ имеет определенное распределение вероятностей на борелевских множествах пространства X. Это не имеет места в общем несепарабельном пространстве X; пример можно найти в работе Недома [1].

Иногда оказывается полезным следующий критерий. Пусть $x_n(\omega)$ — последовательность случайных элементов из X, такая, что $x_n(\omega) \to x(\omega) \in X$ для любого ω. Тогда $x(\omega)$ есть случайный элемент. Бо́льшую информацию можно найти в работе Ганша [1].

Композиция $P * Q$ непрерывна по P и Q. Действительно, предположим, что имеет место слабая сходимость $P_n \to P$, $Q_n \to Q$. Тогда для любой ограниченной и непрерывной функции $f(x)$ при $R_n = P_n * Q_n$ имеем

$$\int_X f(x) R_n(dx) = \int_{x \in X} P_n(dx) \int_{y \in X} f(x+y) Q_n(dy) \to$$
$$\to \int_{x \in X} P(dx) \int_{y \in X} f(x+y) Q(dy) = \int_X f(x) R(dx)$$

(модификации доказательства в замечаниях $2.2._2$, см. также теорему 6.2.2).

6.1.$_2$. Для конечномерных евклидовых пространств это было показано уже в работе Крамера и Вольда [1].

6.1.$_3$. Доказательства свойств 1)—5) операции образования среднего значения (в терминах интеграла Петтиса) можно найти, например, в книге Хилле и Филлипса [1], стр. 91—95.

Заметим, что для сепарабельных пространств, которые мы рассматриваем, интеграл Петтиса эквивалентен интегралу Бохнера и $\mathbf{E}\,\|x\| < \infty$ является не только достаточным, но и необходимым условием существования $\mathbf{E}x$.

Некоторые авторы вводят понятие математического ожидания или среднего значения в более общих пространствах. Если X — метрическое пространство с метрикой $\varrho(x, y)$, то говорят, что случайный элемент x, принимающий значения из X, имеет среднее значение $m \in X$, если

$$\mathbf{E}\varrho^2(x, m) = \inf_{a \in X} \mathbf{E}\varrho^2(x, a).$$

Это определение принадлежит Фреше [1]. Аналогичное определение принадлежит Доссу [1]: случайный элемент x имеет среднее значение $m \in X$, если

$$\varrho(m, a) \leqslant \mathbf{E}\varrho(x, a)$$

для каждого $a \in X$. Так как эти определения выходят за рамки нашего исследования, то они не будут изучаться далее; заинтересованный читатель может посмотреть цитированные выше работы и работу Урбаника [3], в которой изучается интеграл Досса для случая коммутативной группы X с инвариантной метрикой ϱ.

6.1.$_4$. Эта метрика, на действительной прямой введенная Леви, изучалась в рассматриваемой нами ситуации Прохоровым [1], у которого можно найти доказательство утверждения относительно метризации.

6.2.$_1$. (См. Халмош [1], стр. 45.) Для несепарабельного пространства эта ситуация еще полностью не исследована. (См. также Дримл [1].)

6.2.$_2$. (См. Прохоров [1].) На действительной прямой компактность вероятностных мер P_n эквивалентна равностепенной непрерывности $\hat{P}_n$ при $x^* = 0$. Стоит отметить, что не существует локально выпуклой топологии в гильбертовом пространстве l_2, для которой это утверждение имело бы место (см. Прохоров и Сазонов [1]).

6.2.$_3$. Теорема 6.2.4 принадлежит Колмогорову [3]. Некоторые интересные обобщения даны Прохоровым [2].

Предполагалось, что имеет смысл работать со слабыми распределениями: если $x^*(x)$ имеют вполне определенные и согласованные распределения вероятностей для всех $x^* \in X$, то эта система распределений вероятностей называется слабым распределением на X. Нетрудно видеть, что функция $\varphi(x^*)$ определяет слабое распределение тогда и только тогда, когда она является положительно определенной, $\varphi(0) = 1$ и $\varphi(x^*)$ непрерывна вдоль любого луча cx^*, $-\infty < c < \infty$. В настоящее время польза слабых распределений представляется сомнительной.

6.3. Нормальные распределения могут быть эквивалентным образом определены так: $P \in \mathscr{P}(X)$ называется нормальным распределением вероятностей, если $x^*(x)$ — нормальное распределение на R^1 для каждого $x^* \in X^*$.

6.3.$_1$. Это рассуждение основано на следующей теореме Крамера [1]: если x, y, z — действительные случайные величины, x и y независимы, а $z = x + y$ нормальна, то x и y также имеют нормальные распределения.

6.3.$_2$. В доказательстве теоремы 6.3.6 нужно только предположить, что

$$\lim_{N\to\infty} \sup_n \sum_{N+1}^{\infty} s_\nu^n = 0;$$

на самом деле это является также необходимым условием (см. Прохоров [1]).

6.4. Существует много различных вариантов закона больших чисел в банаховом пространстве; некоторые из них принадлежат Мурье [1]. В отдельных теоремах независимость заменяется стационарностью. Бек [1] рассматривал последовательность независимых случайных элементов x_n, таких, что $\mathbf{E}x_n = 0$, $\mathbf{E}\,\|x_n\|^2 \leqslant M < \infty$, и показал, что в этом случае усиленный закон больших чисел имеет место при предположении, что X — равномерно выпуклое банахово пространство.

Можно (и это нетрудно) заменить предположение, что x_ν независимы и одинаково распределены, условием, что они образуют стационарную последовательность. Аналогично можно изучать стационарные процессы с непрерывным параметром, принимающие значения в банаховом пространстве, и соответствующие эргодические теоремы.

6.5. Эти формы центральной предельной теоремы принадлежат Мурье [1] и Форте и Мурье [2]. X называется *G-пространством*, если существуют такая постоянная K и такое отображение g: $X \to X^*$, $g(x) \in X^*$, что

$$1)\quad \|g(x)\| = \|x\|,$$

$$2)\quad (g(x),\ x) = \|x\|^2,$$

$$3)\quad \|g(x) - g(y)\| \leqslant K\,\|x - y\|.$$

Среди G-пространств содержатся L_p-пространства при $p \geqslant 2$ и, в частности, гильбертово пространство.

6.6. Стохастические распределения Шварца были введены Гельфандом [1] и К. Ито [3]. Дальнейшие результаты были получены Ульрихом [1], Урбаником [3] и др.

Гельфанд [1] использует следующее несколько более широкое определение стохастического распределения Шварца: $d(\varphi)$ должно быть обычной случайной величиной для любого φ и аналогично для любого вектора $d(\varphi_1), d(\varphi_2), \ldots, d(\varphi_n)$; d должно быть линейным функционалом: $d(c_1\varphi_1 + c_2\varphi_2) = c_1 d(\varphi_1) + c_2 d(\varphi_2)$; для любого натурального числа n и $\varphi_{i\nu}$, таких, что $\lim\limits_{\nu\to\infty} \varphi_{i\nu} = \varphi_i$, $i = 1, 2, \ldots, n$; конечномерное распределение величин $d(\varphi_{1\nu}), d(\varphi_{2\nu}), \ldots, d(\varphi_{n\nu})$ должно стремиться к распределениям величин $d(\varphi_1), d(\varphi_2), \ldots, d(\varphi_n)$. Гельфанд вводит характеристический функционал

$$\hat{P}(\varphi) = \mathbf{E}\exp[id(\varphi)],\quad \varphi \in \Phi.$$

Как обычно, это — положительно определенная эрмитова функция, удовлетворяющая условию $\hat{P}\ (0) = 1$. Она непрерывна, так как если $\varphi_n \to \varphi$, то распределения обычных случайных величин $d(\varphi_n)$ стремятся к распределению $d(\varphi)$, так что $\hat{P}(\varphi_n) \to \hat{P}(\varphi)$. Более интересно то, что верно и обратное. Прежде всего ясно, что если $f(\varphi)$ — положительно определенная эрмитова функция и $f(0) = 1$, то можно определить согласованные конечномерные распределения вероятностей обычных случайных величин, обозначаемых $x_1 = d\ (\varphi_1),\ x_2 = d\ (\varphi_2),\ \ldots, x_n = d\ (\varphi_n)$, требованием

$$\mathbf{E} \exp\left(i \sum_1^n t_\nu x_\nu\right) = f\left(\sum_1^n t_\nu \varphi_\nu\right).$$

Пусть $x = d\ (\sum_1^n c_\nu \varphi_\nu)$ — случайная величина, соответствующая линейной комбинации $\sum_1^n c_\nu \varphi_\nu$. Тогда

$$\mathbf{E} \exp\left[iu\left(x - \sum_1^n c_\nu x_\nu\right)\right] = f(0) = 1,$$

так что $P\ (x - \sum_1^n c_\nu x_\nu = 0) = 1$ и $d\ (\sum_1^n c_\nu \varphi_\nu) = \sum_1^n c_\nu d\ (\varphi_\nu)$ почти наверное, что и требовалось. Наконец, если $\{\varphi_{1\nu}\}$ такова, как сказано выше, то только что введенные распределения случайных величин $d\ (\varphi_{1\nu}),\ d\ (\varphi_{2\nu}),\ \ldots,\ d\ (\varphi_{n\nu})$ сходятся к распределениям величин $d\ (\varphi_1),\ d\ (\varphi_2),\ \ldots,\ d\ (\varphi_n)$ при $\nu \to \infty$. Следовательно, положительно определенная функция $f(\varphi)$ соответствует стохастическому распределению Шварца, определенному по Гельфанду.

Сравните это понятие с понятием слабого распределения в банаховом пространстве (см. замечания $6.2._3$.).

$6.6._1$. Теорема 6.6.1 принадлежит Ульриху, который доказал ее методом аппроксимации. Существует последовательность $\varphi_1,\ \varphi_2, \ldots$ в Φ, обладающая следующими свойствами. Любая $\varphi \in \Phi$ может быть аппроксимирована подпоследовательностью $\varphi_{n_1},\ \varphi_{n_2}, \ldots$. Но сходимость последовательности

$$\frac{1}{n}\sum_{1}^{n} d_{\nu}(\varphi_i)$$

является непосредственным следствием классического закона больших чисел. Используя равностепенную непрерывность, можно убедиться, что с вероятностью единица

$$\frac{1}{n}\sum_{1}^{n} d_{\nu}(\varphi) \to m(\varphi)$$

для всех $\varphi \in \Phi$.

6.7. Более подробное изложение этих вопросов можно найти в работе Форте и Мурье [1]. Они изучают сходимость эмпирических функций распределения (некоторых случайных вероятностных мер) в общем метрическом пространстве. Их результаты не являются непосредственными следствиями рассмотренных вариантов закона больших чисел, но требуют отдельного вывода.

Другим приложением являются пространства Орлича. Пусть $v = \varphi(u)$ и $u = \psi(v)$ — две взаимно обратные функции, такие, что 1) $\varphi(0) = \psi(0) = 0$ и 2) φ и ψ не убывают. Введем функции

$$\Phi(u) = \int_0^u \varphi(x)\,dx, \quad \Psi(v) = \int_0^v \psi(x)\,dx.$$

На измеримом пространстве $(A, \mathcal{A}, m)$ рассмотрим $\mathcal{A}$-измеримую комплексную функцию $f(a)$, такую, что

$$\sup \int_A |f(a)\,g(a)|\,m(da) < \infty,$$

где верхняя грань берется по всем g таким, что

$$\int_A \Psi[|g(a)|]\,m(da) \leqslant 1.$$

Это множество функций L_Φ называется пространством Орлича. Оно может быть превращено в банахово пространство (см. Заанен [1], стр. 101). Аналогичным образом вводится L_Ψ. Это обобщение пространств L_p. Теперь можно рассмотреть случайные элементы, принимающие значения из L_Φ, и применить общие результаты гл. 6 (см. Бхаруча-Рид [1], [2]).

7.1.$_1$. Для того чтобы убедиться, что $P_x \circ P_y$ вполне определено, заметим, что случайные элементы $x(\omega)$ и $y(\omega)$ могут быть аппроксимированы (сильно) случайными элементами $x_n(\omega)$ и $y_n(\omega)$, принимающими только счетное множество значений. Произведение $x_n(\omega)\cdot y_n(\omega)$ определяет вероятностную меру (также счетнозначную). Так как операция умножения непрерывна, то $x_n(\omega)\cdot y_n(\omega) \to x(\omega)\cdot y(\omega)$, и (см. замечания 6.1.$_1$) отсюда следует, что $x(\omega)\cdot y(\omega)$ есть случайный элемент с вполне определенной вероятностной мерой. Имеем далее

$$\int\limits_{X\times X} f(xy)\, P_x(dx)\, P_y(dy) = \int\limits_X f(z)\, R(dz),$$

где $R = P_x \circ P_y$ для любой ограниченной и непрерывной функции $f(x)$ на X.

Предположим, что имеет место слабая сходимость $P_n \to P$, $Q_n \to Q$. Тогда, как и раньше, можно убедиться в том, что имеет место слабая сходимость $P_n \circ Q_n \to P \circ Q$.

7.1.$_2$. Это выражение может напомнить читателю понятие мультипликативного интеграла. В литературе существует множество определений таких интегралов, но ни одно из них не кажется подходящим для наших целей.

7.1.$_3$. Интересно отметить, что теорема 7.1.5 является частным случаем теоремы 7.1.3. Действительно, условие C_1' в этом случае выполняется, так как

$$\sum_{\nu=1}^{n} \mathbf{E}\,\|y_{n\nu}\| = \frac{1}{n}\sum_{\nu=1}^{n} \mathbf{E}\,\|y_\nu\| = \mathbf{E}\,\|y_1\| < \infty.$$

Далее, для любого $c \in (0,\ 1)$ почти наверное имеет место сильная сходимость

$$\sum_{\nu=1}^{[cn]} y_{n\nu} = \frac{1}{n}\sum_{\nu=1}^{[cn]} y_\nu \to c\mathbf{E}y_1,$$

что следует из закона больших чисел в банаховом пространстве. Но тогда теорема 7.1.3 утверждает, что произведение γ_n должно стремиться по вероятности к случайному элементу $\exp m$, который здесь сводится к постоянному элементу.

7.2.$_1$. Относительно этого и связанных с ним утверждений в тексте см. Хилле и Филлипс [1], стр. 132.

7.2.$_2$. В матричном случае этот результат получен Ферстенбергом и Кестеном [1], причем они показали также, что в этом случае $\frac{1}{n}\log \| x_1 x_2 \ldots \gamma_n \|$ действительно сходится почти наверное.

7.2.$_3$. См. Хилле и Филлипс [1], стр. 138.

7.2.$_4$. См., например, Халмош [1], стр. 114.

7.3. Свойства регулярности стохастических операторов, резольвент и т. д. изучались разными авторами, из которых мы укажем Бхаруча-Рида [1], [2], Ганша [2], Шпачека [1]. Теорема 7.3.1 принадлежит Ганшу.

7.3.$_1$. См. Банах [1].

7.3.$_2$. Идея случайных эргодических теорем восходит к Уламу и фон Нейману; вариант такой теоремы, данный в теореме 7.3.3, принадлежит Какутани [1]. Интересно следующее ее применение (см. Роббинс [1]). Пусть G — компактная группа, и пусть случайные преобразования x_ω определяются как преобразования gh группы G в G, где h имеет распределение вероятностей на G. Тогда x_ω сохраняет меру Хаара, и из случайной эргодической теоремы вытекает, что

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\nu=1}^{n} f(h_1 h_2 \ldots h_\nu)$$

существует почти наверное. Далее, можно показать, что в этом частном случае данный предел почти наверное является константой.

7.4.$_1$. См. Наймарк [1], стр. 164.

7.4.$_2$. См. Наймарк [1], стр. 182.

7.5. Беллман [1] использовал произведения Кронекера для изучения произведений стохастических матриц, в частности для получения моментов элементов произведения матриц.

7.5.1. Матричнозначные мультипликативные процессы были рассмотрены впервые Гренандером [3].

7.5.2. Утверждения этого раздела были доказаны в работе Ферстенберга и Кестена [1].

7.5.3. См. Андерсон [1], гл. 13, и Найквист, Райс и Риордан [1].

Асимптотический спектр стохастических симметричных матриц был получен Уигнером [1]. Нам хотелось бы добавить следующее замечание. Для того чтобы непосредственно проверить сходимость функции мощности, упомянутой в тексте, можно было бы поступить следующим образом. Введем характеристические функции

$$\varphi_n(z) = \int_{-\infty}^{\infty} \exp(i\lambda z)\, dF_n(\lambda) = \frac{1}{n} \sum_{\nu=1}^{n} \exp(i\lambda_\nu z).$$

Имеем

$$\varphi_n(z) = \sum_{p=0}^{2k-1} \mu_p^{(n)} \frac{(iz)^p}{p!} + O\left(\mu_{2k}^{(n)} \frac{z^{2k}}{(2k)!} \right),$$

так что при $n \to \infty$ имеем

$$\mathbf{E}\varphi_n(z) \to \sum_{p=0}^{2k-1} c_{2p} \frac{(iz)^p}{p!} + O\left(c_{2k} \frac{z^{2k}}{(2k)!} \right).$$

Остаточный член может быть при фиксированном z сделан сколь угодно малым, если взять достаточно большое k. Следовательно,

$$\mathbf{E}\varphi_n(z) \to \varphi(z) = \frac{1}{\pi} \int_{-2}^{2} \exp(i\lambda z) \sqrt{1 - \frac{u^2}{4}}\, du$$

для любого действительного значения z. Но если ввести $\mathbf{E}\, F_n(\lambda) = G_n(\lambda)$, то это эквивалентно

$$\int_{-\infty}^{\infty} \exp(iz\lambda)\, dG_n(\lambda) \to \varphi(z),$$

что возможно только тогда, когда

$$G_n(\beta) - G_n(\alpha) \to \frac{1}{\pi} \int_{\alpha}^{\beta} \sqrt{1 - \frac{u^2}{4}}\, du,$$

так что распределения G_n стремятся к предельному распределению с функцией мощности

$$s(u) = \frac{1}{\pi} \sqrt{1 - \frac{u^2}{4}}\,.$$

Хотя результаты, обсуждаемые в разд. 7.5.3, имеют и независимый интерес, представляется, что общая теория спектров стохастических операторов должна использовать более существенные алгебраические свойства операторов. См. также F. J. Dyson, A Brownian-motion model for the eigen-values of a random matrix, *J. Math. and Phys.*, **3** (1962).

7.5.4. См. Уолл [1], стр. 42. Об изучении стохастических интегральных уравнений см. Бхаруча-Рид [1], [2].

ЛИТЕРАТУРА[1])

Андерсон Т.
[1] Введение в многомерный статистический анализ, Физматгиз, М., 1963.
Ахиезер Н. И. и Глазман И. М.
[1] Теория линейных операторов в гильбертовом пространстве, ГИТТЛ, М., 1950.
Банах (Banach S.)
[1] Théorie des opérations linéaires, Warszawa, 1932. (На украинском языке: «Курс функціонального аналізу, Київ, 1948.)
Бёге (Böge W.)
[1] Über die Charakterisierung unendlich teilbarer Wahrscheinlichkeitsverteilungen, *J. reine und angew. Math.*, **201** (1959).
Бек (Beck A.)
[1] Une loi forte des grands nombres dans des espace de Banach uniformément convexes, *Ann. Inst. Henri Poincaré*, **16** (1958).
Беллман (Bellman R.)
[1] Limit theorems for non-commutative operations I, *Duke Math. J.*, **21** (1954).
Бергстрём (Bergström H.)
[1] Konvergenzsätze über unendliche Produkte in abstrakten Halbgruppen, *Abhandl. Math. Seminar Univ. Hamburg*, **23** (1959).
Бохнер (Bochner S.)
[1] Harmonic analysis and the theory of probability, Berkeley, 1955.
Бурбаки Н.
[1] Общая топология. Основные структуры (гл. III), М., Физматгиз, 1958.
Бхаруча-Рид (Bharucha-Reid A. T.)
[1] On random elements on Orlicz spaces, *Bull. Acad. Pol. Sci.*, **4** (1956).

---

[1]) Ссылки на оригинальные работы на иностранных языках и переводы русских работ на эти языки заменены, соответственно, ссылками на переводы на русский язык (в тех случаях, когда они существуют) и на оригинальные работы. Работы перечисляются в русском алфавитном порядке фамилий их авторов.— *Прим. перев.*

[2] Über die Konvergenz der Folgen von verallgemeinerten zufälligen Grössen in Orlicz'schen Räumen, *Bull. Acad. Pol. Sci.*, **7** (1959).
[3] On random solutions of Fredholms integral equations, *Bull. Amer. Math., Soc.*, **66** (1960).
[4] On random solutions of integral equations in Banach spaces, Trans. Second Prague Conf. Inform. Theory, Stat. Decision Functions, Random Processes, Prague, 1960.

Варадараджан (Varadarajan V. S.)
[1] Weak convergence of measures on separable metric spaces, Sankhyā, **19** (1958).

Вейль А.
[1] Интегрирование в топологических группах и его применения ИЛ, М., 1950.

Вен (Wehn D. F.)
[1] Probabilities on Lie groups, *Proc. Nat. Acad. Sci.*, **48**, № 5 (1962).
[2] Limit distributions on Lie groups (в печати).

Воробьев Н. Н.
[1] Сложение независимых случайных величин на конечных абелевых группах. *Матем. сб.*, **34** (1954), 89—126.

Вулл (Woll J. W.)
[1] Homogeneous stochastic processes, *Pacif. J. Math.*, **9** (1959).

Ганш (Hanš O.)
[1] Generalized random variables, Trans. First Prague Conf. Inform. Theory, Stat. Decision Functions, Random Processes. Prague, 1957.
[2] Random fixed point theorems, Trans. First Prague Conf. Inform. Theory. Stat. Decision Functions, Random Processes, Prague, 1957.

Гельфанд И. М.
[1] Обобщенные случайные процессы. *ДАН СССР*, **100** (1955), 853—856

Гельфанд И. М. и Наймарк М. А.
[1] Унитарные представления классических групп, *Труды МИАН*, **36** (1950).

Герценштейн М. Е. и Васильев В. Б.
[1] Волноводы со случайными неоднородностями и броуновское движение в плоскости Лобачевского, *Теория вероятн. и ее примен.*, 4 (1959), 424—432.

Гетур (Getoor R. K.)
[1] Infinitely divisible probabilities on the hyperbolic plane, *Pacif. J. Math.*, **11** (1961).

Главка (Hlawka E.)
[1] Statistik auf kompakten Gruppen, I. *Österr. Akad. Wiss.*, 1959.

Гликсберг (Glicksberg I.)
[1] Convolution semigroups of measures, *Pacif. J. Math.*, **9** (1959).

Гнеденко Б. В. и Колмогоров А. Н.
[1] Предельные распределения для сумм независимых случайных величин, ГИТТЛ, М., 1949.

Годман (Godement R.)
[1] Les fonctions de type positif et la théorie des groupes, *Trans. Amer. Math. Soc.*, **63** (1948).
[2] Sur la théorie des représentations unitaires, *Ann. Math.*, **53** (1951).
Горман (Gorman C. D.)
[1] Brownian motion of rotation, *Trans. Amer. Math. Soc.*, **94** (1960).
Гренандер (Grenander U.)
[1] Some non-linear problems in probability theory, Harald Cramér volume, New York, 1959.
[2] Stochastic groups. Background. General discussion. Remarks on limit theorems, *Arkiv mat.*, **4** (1960).
[3] Stochastic groups. A particular stochastic group. Some stochastic algebras, *Arkiv mat.*, **4** (1960).
[4] Stochastic groups. Fourier analysis of probability distributions on locally compact groups, *Arkiv. mat.*, **5** (1961).
[5] Stochastic groups and related structures, *Proc. Fourth Berkeley Symp. on Math. Stat. and Prob.*, 1961.
Гренандер и Розенблат (Grenander U., Rosenblatt M.)
[1] Statistical analysis of stationary time series, New York, 1957.
Дайсон (Dyson F. J.)
[1] The dynamics of a disordered linear chain, *Phys. Rev.*, **101** (1956).
Данфорд Н. и Шварц Дж. Т.
[1] Линейные операторы. Общая теория, ИЛ, М., 1962.
Дворецкий и Вольфовиц (Dvoretzky A., Wolfowitz J.)
[1] Sums of random integers reduced modulo *m*, *Duke Math. J.*, **18** (1951).
Донскер (Donsker M. D.)
[1] An invariance principle for certain probability limit theorems *Mem. Amer. Math. Soc.*, **6** (1951).
Досс (Doss S.)
[1] Sur la moyenne d'un élément aléatoire dans un espace distancié, *Bull. sci. math.*, **73** (1949).
[2] Sur la convergence stochastique dans les espaces uniformes, *Ann. scient. École norm. supér*, **71** (1954).
Дримл (Driml M.)
[1] Convergence of compact measures on metric spaces, Trans. Second Prague Conference. Inform. theory. Stat. decision functions, Random processes, Prague, 1960.
Дуб Дж. Л.
[1] Вероятностные процессы, ИЛ., М., 1956.
Дьедонне (Dieudonné J.)
[1] Sur le produit de composition, *Compositio math.*, **12** (1954).
[2] Sur le produit de composition II, *J. math. pures et appl.*, **39** (1960).
[3] Sur une propriété des groupes libres, *J. reine und angew. Math.*, **204** (1960).

Заанен (Zaanen A. C.)
[1] Linear analysis, Amsterdam, 1953.
Ито К. (Itô K.)
[1] Stochastic differential equations in a differentiable manifold, *Nagoya Math. J.*, **1** (1950).
[2] Brownian motions in a Lie group, *Proc. Japan, Acad.*, **26** (1950).
[3] Stationary random distributions, *Mem. Coll. Sci. Univ. Kyoto*, **28** (1954).
Ито С. (Itô S.)
[1] Fundamental solutions of parabolic differential equations and boundary value problems, *Japan. J. Math.*, **27** (1957).
[2] Brownian motions in a topological group and its covering group. *Rend. Circolo. mat. Palermo*, **1** (1952).
Иосида и Какутани (Yosida K., Kakutani S.)
[1] Operator theoretical treatment of Markov processes and mean ergodic theorem, *Ann. Math.*, **42** (1941).
Йосидзава (Yoshizawa H.)
[1] Some remarks on unitary representations of the free group: *Osaka Math. J.*, **3** (1951).
Кавада и Ито К. (Kawada Y., Itô K.)
[1] On the probability distribution on a compact group I. *Proc. Phys.-Math. Soc. Japan*, **22** (1940).
Какутани (Kakutani S.)
[1] Random ergodic theorems and Markov processes with a stable distribution, Proc. Second Berkeley Symp., Math. Stat. Prob., Berkeley, 1951.
Какэхаси (Kakehashi T.)
[1] Stationary periodic distributions. *J. Osaka Inst. Sci. Technology*, **1** (1949).
Каллианпур (Kallianpur G. B.)
[1] The topology of weak convergence of probability measures *J. Mat. and Mech.*, **10** (1961).
Карналь (Carnal H.)
[1] Distributions aléatoires indéfiniment divisible sur les groupes compacts, *C. r. Acad. sci.*, **255** (1962).
Карпелевич Ф. И., Тутубалин В. Н., Шур М. Г.
[1] Предельные теоремы для композиций распределений в плоскости и пространстве Лобачевского, *Теория вероятн. и ее примен.*, **4** (1959).
Кестен (Kesten H.)
[1] Symmetric random walks on groups, *Trans. Amer. Math. Soc.*, **92** (1959).
[2] Full Banach mean values on countable groups, *Math. Scand.*, **7**, (1959).
Клосс Б. М.
[1] О вероятностных распределениях на бикомпактных топологических группах, *Теория вероятн. и ее примен.*, **4** (1959), 255—290.
[2] Предельные распределения для сумм независимых случай-

ных величин, принимающих значения из бикомпактной группы, *ДАН СССР*, **109** (1953), 453—455.

[3] Предельные распределения на бикомпактных абелевых группах, *Теория вероятн. и ее примен.*, **6** (1961), 392—421.

Коллинз (Collins H. S.)

[1] Primitive idempotents in the semigroup of measures, *Duke Math. J.*, **27** (1960).

[2] Convergence of convolution iterates of measures, *Duke Math. J.*, **29** (1962).

Колмогоров А. Н.

[1] Основные понятия теории вероятностей, ОНТИ, М.—Л.,1936.

[2] La transformation de Laplace dans les espaces linéaires, *Compt. rend.*, **200** (1935).

[3] Замечание о работах Р. А. Минлоса и В. В. Сазонова, *Теория вероятн. и ее примен.*, **4** (1959), 237—239. (Резюме докладов семинара кафедры теории вероятностей МГУ.)

Кон (Cohn P. M.)

[1] Lie groups, Cambridge, 1957.

Кох и Уоллес (Koch R. J., Wallace A. D.)

[1] Maximal ideals in compact semigroups, *Duke Math. J.*, **21** (1954).

Крамер Г. (Cramér H.)

[1] Über eine Eigenschaft der normalen Verteilungsfunktion, *Math. Z.*, **41** (1936).

[2] Математические методы статистики, ИЛ, М., 1948.

Крамер и Волд (Cramér H., Wold H.)

[1] Some theorems on distribution functions, *J. London, Math. Soc.*, **11** (1936).

Леви (Lévy P.)

[1] L'addition des variables aléatoires définies sur une circonférence, *Bull. Math. France*, **67** (1939).

Ле Кам (Le Cam L.)

[1] Un instrument d'étude des fonctions aléatoires: la fonctionelle caractéristique, *Compt. rend.*, **224** (1947).

[2] Convergence in distribution of stochastic processes, Publ. Statistics, Univ. California Press, Berkeley, 1957.

Лойнес (Loynes R.)

[1] Fourier transforms and probability theory on a non-commutative locally compact groups, *Arkiv mat.*, **5** (1963), 37—42.

Лоэв М.

[1] Теория вероятностей, ИЛ, М., 1962.

Люмис Л.

[1] Введение в абстрактный гармонический анализ, ИЛ М. 1956.

Мак-Кин (McKean H. P.)

[1] Brownian motions on the 3-dimensional rotation group, *Mem. Coll. Sci. Univ. Kyoto*, **33** (1960).

Маутнер (Mautner F. I.)

[1] Unitary representations of locally compact groups II, *Ann. Math.*, **52** (1950).

[2] Note on the Fourier inversion formula on groups, *Trans. Amer. Math. Soc.*, **78** (1955).

Мурье (Mourier E.)

[1] Eléments aléatoires dans un espace de Banach, *Ann. Inst. Henri Poincaré*, **13** (1953).

Найквист, Райс и Риордан (Nyquist H., Rice S. O., Riordan J.)

[1] The distribution of random determinants, *Quart. Appl. Math.*, **12** (1954).

Наймарк М. А.

[1] Нормированные кольца, ГИТТЛ, М., 1956.

Недома (Nedoma J).

[1] Note on generalized random variables, Trans. First Praque Conf. Inform. Theory, Stat. Decision. Functions, Random Processes, Prague, 1957.

Нумакура (Numakura K.)

[1] On bicompact semigroups, *Math. J. Okayama Univ.*, **1** (1952).

Пакшираджан (Pakshirajan R. P.)

[1] Regular measures and stochastic processes in topological groups (miméographed report), Univ. Oregon, 1959.

[2] A property of regular measures in locally compact Hausdorff spaces, *Proc. Amer. Math. Soc.*, **12** (1961).

Партасарати, Ранга Рао и Варадхан (Parthasarathy K. R., Ranga Rao R., Varadhan S. R. S.)

[1] On the category of indecomposable distributions on topological groups, *Trans. Amer. Math. Soc.*, **102** (1962).

[2] Распределения вероятностей на локально компактных абелевых группах, *Математика*, **9 : 2** (1965).

Перрэн (Perrin F.)

[1] Étude mathématique du mouvement Brownien de rotation, *Ann. scient. École norm. supér.*, 1928.

Понтрягин Л. С.

[1] Топологические группы, изд. 2-е, ГИТТЛ, М., 1954.

Прекопа А., Реньи А. и Урбаник К.

[1] О предельном распределении для сумм независимых случайных величин на бикомпактных коммутативных топологических группах, *Acta math. Acad. sci. hung.*, **7** (1956), 11—16.

Прохоров Ю. В.

[1] Сходимость случайных процессов и предельные теоремы теории вероятностей, *Теория вероятн. и ее примен.*, **1** (1956), 177—238.

[2] The method of characteristic functionals, *Proc. fourth. Berkeley Symp. on Math. Stat. and Prob.*, 1961.

Прохоров Ю. В. и Сазонов В. В.

[1] Некоторые результаты, связанные с теоремой Бохнера, *Теория вероятн. и ее примен.*, **6** (1961), 87—93.

Ривкинд Я. И.

[1] Предельная теорема теории вероятностей на компактных топологических группах., *Уч. зап. Гродненск. пед. ин-та*, 1955.

Рисс Ф. и Секефальви-Надь Б.

[1] Лекции по функциональному анализу. ИЛ., М., 1954.

Роббинс (Robbins H.)
[1] On the equidistribution of sums of independent random variables, *Proc. Amer. Math. Soc.*, **4** (1953).
Розенблат (Rosenblatt M.)
[1] Limits of convolution sequences of measures on a compact topological semigroup, *J. Math. and Mech.*, **9** (1960).
Рудин (Rudin W.)
[1] Measure algebras on Abelian groups, *Bull. Amer. Math. Soc.*, **65** (1959).
[2] Idempotent measures on Abelian groups, *Pacif. J. Math.*, **9** (1959), 195—209.
Стромберг (Stromberg K.)
[1] Probabilities on a compact group, *Trans. Amer. Math. Soc.*, **94** (1960).
[2] A note on convolution of semigroups, *Math. Scand.*, 7 (1959).
Такенути (Takenouchi O.)
[1] Sur une classe de fonctions continues de type positif sur un groupe localement compact, *Math. J. Okayama Univ.*, **4** (1955).
Уигнер (Wigner E. P.)
[1] On the distribution of the roots of certain symmetric matrices, *Ann. Math.*, **67** (1958).
Ульрих (Ullrich M.)
[1] Some theorems on random Schwartz distributions, Trans. First Prague Conf. Inform. theory, Stat. Decision functions, Random Processes, Prague, 1957.
Уолл (Wall H. S.)
[1] Analytic theory of continued fractions, New York, 1948.
Уоллес (Wallace A. D.)
[1] The Rees-Suschkewitsch structure theorem for compact simple semigroups, *Proc. Nat. Acad. Sci. U.S.A.*, **42** (1956).
Урбаник (Urbanik K.)
[1] On the limiting probability distribution on a compact topological group, *Fundam. Math.*, **44** (1957).
[2] Poisson distributions on compact Abelian topological groups, *Colloq. math.*, **6** (1958).
[3] Remarks on the Doss integral, *Colloq. math.*, **5** (1957).
[4] Generalized stochastic processes with independent values, *Proc. Fourth. Berkeley Symp. on Math. Stat. and Prob.*, 1961.
Уэндел (Wendel J. G.)
[1] Haar measure and the semigroup of measures on a compact group, *Proc. Amer. Math. Soc.*, **5** (1954).
Ферстенберг и Кестен (Furstenberg H., Kesten H.)
[1] Products on random matrices, *Ann. Math. Stat.*, **31** (1960).
Форте (Fortet R.)
[1] Normalverteilte Zufallselemente in Banachschen Räumen, Anwendungen auf zufällige Funktionen, Tagung Wahrscheinl. Math. Stat., Berlin, 1954.
Форте и Мурье (Fortet R., Mourier E.)
[1] Convergence de la répartition empirique vers la répartition théorique, *Ann. scient. Ecole. norm. supér.*, **60** (1953).

[2] Les fonctions aléatoires comme éléments aléatoires dans les espace de Banach, *Studia math.*, **15** (1955).

Фреше (Fréchet M.)

[1] Les eléments aleatoires de nature quelconque dans un espace distancié, *Ann. Inst. H. Poincaré*, **10** (1947).

[2] Abstrakte Zufallselemente, Tagung Wahrscheinl. Math. Stat., Berlin, 1954.

Халмош П.

[1] Теория меры, ИЛ, М., 1953.

Хант (Hunt G. A.)

[1] Semigroups of measures on Lie groups, *Trans. Amer. Math. Soc.*, **81** (1956).

Хебл и Розенблат (Heble M., Rosenblatt M.)

[1] Idempotent measures on a compact topological semigroup, *Proc. Amer. Math. Soc.*, **14** (1963).

Хилле Э. и Филлипс Р. С.

[1] Функциональный анализ и полугруппы, ИЛ, М., 1962.

Хинчин А. Я.

[1] Асимптотические законы теории вероятностей, М.— Л., 1936.

Хьюит (Hewitt E.)

[1] A survey of abstract harmonic analysis, Surveys in Applied Math. IV, New York, 1958.

Хьюит и Рубин (Hewitt E., Rubin H.)

[1] The maximum value of a Fourier-Stieltjes Transform, *Math. Scand.*, **3** (1955).

Хьюит и Цуккерман (Hewitt E., Zuckerman H. S.)

[1] Arithmetic and limit theorems for a class of random variables, *Duke Math. J.*, **22** (1955).

Циглер (Cigler J.)

[1] Folgen normierter Masse auf kompakten Gruppen, *Z. Wahrscheinlichkeitstheorie und verw. Gebiete*, **1** (1962).

Шварц (Schwarz S.)

[1] On the structure of the semigroup of measures on a finite semigroup, *Czechosl. Math. J.*, **7** (1957).

Шпачек (Špaček A.)

[1] Prolongement des transformations aléatoires, Proc. First Prague Conf. Inform. Theory, Stat. Decision Functions, Random Processes, Prague, 1960.

[2] Probability measures in infinite Cartesian products, *Ill. J. Math.*, **4** (1960).

УКАЗАТЕЛЬ

ОГЛАВЛЕНИЕ