

## ПРЕДИСЛОВИЕ

Настоящее издание содержит избранные работы выдающегося математика Николая Николаевича Ченцова (1930–1992). Богатство идей, которыми насыщены публикуемые работы, новые постановки задач, новые мощные методы исследования, часто весьма далекие от традиционно применяемых, умение находить содержательную математическую постановку актуальных прикладных вопросов делают это собрание работ Н.Н. Ченцова интересным для широкого круга математиков и научных работников, занимающихся прикладной математикой.

Работы Н.Н. Ченцова в области прикладной математики отмечены двумя Государственными премиями (в составе коллектива) и орденом Трудового Красного Знамени.

Настоящее издание состоит из пяти разделов и включает значительную часть работ Н.Н. Ченцова, опубликованных в 1956–1991 гг. по теории случайных процессов и полей, по математической статистике, некоммутативной теории вероятностей, теории кубатурных формул и методу Монте-Карло. В разделе V приведены работы, содержащие методы и решения интересных физических задач, в том числе задач излучения и кумуляции энергии.

При отборе статей предпочтение отдавалось тем из них, которые, по мнению редколлегии, наиболее интересны в научном плане, а также тем, которые были первоначально опубликованы в труднодоступных или недоступных для широкого читателя изданиях, например в журнале “Вопросы атомной науки и техники” (ВАНТ). Поэтому остались невключенными в данное издание важные работы, например, такие, как “Решение задач об осесимметричном движении газа с ударной волной” (совместно с К.И. Бабенко, И.М. Гельфандом, М.В. Келдышем, О.В. Локуциевским), которую читатель найдет в книге: М.В. Келдыш “Избранные труды. Математика”, Т. 1 (М.: Наука, 1985. С. 377–420). Работа “Естественная геометрия семейств вероятностных законов” (совместно с Е.А. Морозовой), опубликованная в серии “Современные проблемы математики. Фундаментальные направления” (М.: ВИНТИ, 1991. Т. 83. С. 133–265), представлена только оглавлением, введением, заключением и списком литературы (198 наименований), что дает возможность предварительно ознакомиться с кругом идей нового направления.

Никаким изменениям, исключая исправления мелких погрешностей и опечаток, текст публикуемых работ не подвергался. Для удобства чтения обозначения по возможности приведены к единому виду.

Работы внутри разделов расположены в хронологическом порядке.

В конце книги приведены краткие комментарии к разделам и список работ Н.Н. Ченцова. Работы, включенные в настоящее издание, отмечены “галочкой”.

В список не включены, к сожалению, несколько десятков статей, написанных Н.Н. Ченцовым для энциклопедических изданий, а также мемориальные и критические статьи.

Характеристика научного творчества Н.Н. Ченцова не может быть полной без упоминания о его фундаментальной монографии “Статистические решающие правила и оптимальные выводы” (М.: Наука, 1972), переведенной в 1982 г. Американским математическим обществом. Многолетний опыт работы со школьниками, увлекающимися математикой, нашел отражение в книгах, одним из соавторов которых был Николай Николаевич Ченцов, “Избранные задачи и теоремы элементарной математики” (Т. 1. “Арифметика”, Т. 2. “Планиметрия”, Т. 3. “Стереометрия”). Эти книги стали неотъемлемой частью библиотеки математических кружков. Они, естественно, также не вошли в настоящее издание.

Редколлегия благодарит за большую работу по подготовке издания книги Наталью Николаевну Ченцову, редактора Елену Юрьевну Ходан, а также Николая Николаевича Андреева за компьютерный набор и изготовление оригинал-макета.

Редколлегия выражает глубокую признательность Российскому фонду фундаментальных исследований, финансовая поддержка которого позволила осуществить настоящее издание.

*Редколлегия*

## НИКОЛАЙ НИКОЛАЕВИЧ ЧЕНЦОВ \*)

Николай Николаевич Ченцов родился 19 февраля 1930 года в Москве. Его отец — Николай Гаврилович Ченцов (1882–1968) — окончил математическое отделение физико-математического факультета Московского университета в 1904 г. и работал научным сотрудником ЦАГИ (с момента основания института в 1918 г. вплоть до выхода на пенсию в 1958 г.). Ученик и сотрудник Николая Егоровича Жуковского, профессор высшего технического училища, Герой Труда, он занимался теорией упругости, теорией композитов, газовой динамикой. Мать — Екатерина Ивановна Ченцова (урожденная Дорофеева, 1891–1960) — окончила пять курсов медицинского отделения Московских Высших женских курсов и в 1914–1919 гг. работала в различных госпиталях сестрой милосердия, а затем фельдшером. После рождения детей не работала, занимаясь их воспитанием.

Интерес к математике у Коли Ченцова проявился достаточно рано, а судьба свела его со многими замечательными математиками и физиками. Восьмиклассником он пришел в школьный математический кружок при механико-математическом факультете Московского государственного университета им. М.В. Ломоносова, где его первыми учителями стали вначале аспирант А.С. Кронрод и студентка О.А. Ладыженская, а в 10 классе — аспирант Е.Б. Дынкин. На X Московской математической олимпиаде десятиклассник Коля Ченцов получил I премию.

В 1947 г. Н.Н. Ченцов поступил на механико-математический факультет МГУ и очень активно включился в руководство школьным кружком и проведение олимпиад. Он был ответственным секретарем XIV Московской математической олимпиады и секретарем Совета по олимпиадам при ректорате МГУ (одновременно с Ремом Хохловым). Студент Коля Ченцов был одним из авторов трехтомника “Избранные задачи и теоремы элементарной математики”. Особенно много труда вложил он в написание третьего тома “Стереометрия”.

Атмосфера на мехмате располагала к активным занятиям математикой. Он работал в семинаре А.Н. Колмогорова, влияние которого испытывал до конца своих дней, сдавал много специальных курсов. А научная работа Н.Н. Ченцова началась с I курса в семинаре Е.Б. Дынкина, под его руководством он пишет дипломную работу “Асимптотическая теория статистических оценок”.

\*) Настоящий очерк повторяет с небольшими изменениями статью, посвященную памяти Н.Н. Ченцова, которую написали В.И. Арнольд, Н.С. Бахвалов, К.В. Брушлинский, И.М. Гельфанд, Р.Л. Добрушин, А.В. Забродин, И.А. Ибрагимов, С.П. Курдюмов, С.П. Новиков, Д.Е. Охочимский, Ю.В. Прохоров, Ю.Б. Радвогин, Я.Г. Синай, Р.З. Хасьминский, А.С. Холево, А.Н. Ширяев, Т.М. Энеев, опубликованную в журнале “Успехи математических наук” (1993. Т. 48, вып. 2). См. также: Теория вероятностей и ее применение. 1993, Т. 38, вып. 3.

В 1952 г. он с отличием окончил мехмат и получил официальную рекомендацию в аспирантуру МГУ, но по распоряжению И.Г. Петровского был направлен в Расчетное бюро Математического института им. В.А. Стеклова АН СССР. Там в то время проводились вычислительные работы в обеспечение решения атомной проблемы. Руководил этими работами Мстислав Всеволодович Келдыш. По просьбе И.Г. Петровского Н.Н. Ченцову в порядке исключения было разрешено сочетать работу с учебой в заочной аспирантуре МИАН, где он продолжил занятия статистикой у Н.В. Смирнова.

В Расчетном бюро Н.Н. Ченцов был определен в группу, возглавляемую И.М. Гельфандом, с которым он проработал более 30 лет. Научные контакты с И.М. Гельфандом требовали глубокого овладения современной математикой и физикой. Основным содержанием деятельности Н.Н. Ченцова в 1952–1959 гг. стали разработка алгоритмов и проведение расчетов в задачах переноса и рассеяния излучения в сложных системах. Эти расчеты проводились по заданиям Я.Б. Зельдовича, А.Д. Сахарова и их сотрудников. За их выполнение Н.Н. Ченцов в 1956 г. был награжден орденом Трудового Красного Знамени. Последующие результаты этого цикла работ были отмечены присуждением ему в составе авторского коллектива Государственной премии (1972 г.).

Одновременно с этими работами Н.Н. Ченцов совместно с М.В. Келдышем, К.И. Бабенко, Н.А. Дмитриевым и другими сотрудниками уже Отделения прикладной математики МИАН (в которое в 1953 г. объединением с группой А.Н. Тихонова было преобразовано Расчетное бюро) участвовал в пионерской работе по расчету нестационарного газодинамического течения, порождаемого движением осесимметрической ударной волны. В этой работе были разработаны и реализованы оригинальный способ математического описания двумерного течения и вычислительный алгоритм решения подобных задач — “матричная прогонка”.

Научный интерес к проблеме кумуляции сохранился у Н.Н. Ченцова на протяжении всех лет его работы. Вместе с Ю.Н. Бабаевым и другими соавторами им были разработаны оригинальные газодинамические конструкции, реализующие сферическую кумуляцию, и построены приближенные методы, позволяющие проводить предварительный отбор решений, близких к оптимальному, по отбору кумулирующей энергии. Эти исследования являют собой красивое и эффективное сочетание аналитических построений с результатами численных расчетов двумерных нестационарных течений со сложной геометрией.

В то же время Н.Н. Ченцов выполнил в семинаре Е.Б. Дынкина работу “Слабая сходимость случайных процессов с траекториями без разрывов второго рода и так называемый “эвристический” подход к критериям согласия типа Колмогорова–Смирнова” (см. [2]). В этой работе он дал простой критерий, при выполнении которого траектории сепарабельного процесса с вероятностью единица не имеют разрывов второго рода. Этот критерий получил название “критерий Колмогорова–Ченцова”, так как он обобщает критерий Колмогорова непрерывности почти наверное сепарабельного процесса. Из этой работы следует обоснование непараметрических критериев (типа критерия Колмогорова–Смирнова). При полной строгости и ясности подход Ченцова является, по-видимому, простейшим из возможных.

Эти результаты легли в основу кандидатской диссертации Н.Н. Ченцова “Обоснование статистических критериев методами случайных процессов”, защищенной в 1958 г. в МИАН.

Работа “Многопараметрическое броуновское движение Леви и обобщенный белый шум” [4] посвящена конструкции методами интегральной геометрии случайных полей Леви  $\zeta(t)$  — аналогов винеровского процесса, у которого время заменено многомерным параметром, пробегающим все  $n$ -мерное пространство или  $n$ -мерную сферу. Конструкция, связывающая поля Леви с белым шумом на грассманианах, доказывала существование таких полей и сильно упрощала их теорию. Эти две работы принесли Н.Н. Ченцову широкую известность. Следует отметить еще многомерное обобщение броуновского движения — случайное поле Китагавы–Ченцова — и его применение к вычислению предельных распределений некоторых функционалов (распространению принципа Дуба).

Исследования сепарабельных процессов без разрывов второго рода заставили Н.Н. Ченцова изучить теорию меры в тихоновских степенях  $X^T$ , для которых он дал (1958 г.) первые примеры неборелевских подмножеств пространства  $\mathbb{R}^T$ , где  $T$  — отрезок вещественной прямой. Эти результаты были изложены в статье “Неизмеримые подмножества Тихоновского куба” [53].

Практическим выходом его теоретических исследований по теории вероятностей и математической статистике явился большой вклад в развитие и применение методов моделирования, отмеченный в 1979 г. Государственной премией СССР (в коллективе авторов). Его работы в этой области можно условно разбить на три направления.

1. Принципы метода Монте-Карло, способы ускорения вычислений и обработок результатов (совместно с сотрудником и учеником А.С. Фроловым).

2. Задачи теории переноса — расчет критического параметра реактора на быстрых нейтронах, распределения гамма-излучения в воздухе и расчете характеристик нейтронного спектрометра.

3. Кубатурные формулы для кратных и бесконечнократных интегралов. (Теория численных кубатур бесконечнократных интегралов ведет начало с работы И.М. Гельфанда и Н.Н. Ченцова “О численном вычислении континуальных интегралов” [3] и исследований Н.Н. Ченцова “О квадратурных формулах для функции бесконечно большого числа переменных” [14].)

Большое значение в становлении советской школы статистического моделирования имели лекции Н.Н. Ченцова по основам и принципам метода Монте-Карло, прочитанные в 1958–1959 и 1959–1960 учебных годах в МГУ, а также в ряде научных школ тех лет по вычислительной математике.

Вместе с Н.М. Коробовым и Н.С. Бахваловым, Н.Н. Ченцов руководил семинаром в Отделе теории чисел МИАН, где рассматривались вопросы вычисления интегралов высокой размерности и конструирования псевдослучайных чисел.

Занимаясь в 60-е годы проблемой наиболее точного определения плотности случайной величины по наблюдениям, Н.Н. Ченцов пришел к идее поиска естественной геометрии семейств вероятностных законов. Различные геометрические аспекты математической статистики обсуждались тогда А.Н. Колмогоровым, С.Р. Рао, С. Кульбаком и другими исследователями. Итогом изысканий Н.Н. Ченцова явилось создание в 1964 г. красивой своеобразной геометрии ста-

тистических решений с категорией марковских отображений (обобщающей группу движений в обычной геометрии). В терминах этой фундаментальной категории естественно описываются наиболее существенные, глубинные свойства объектов теории статистических решений А. Вальда, Л. Блекуэлла, Ч. Стейна, получает прозрачное объяснение исключительная роль фишеровского информационного тензора как порождающего единственную монотонно инвариантную метрику и экспоненциальных семейств как геодезических относительно некоторой инвариантной аффинной связности. Н.Н. Ченцовым было открыто целое семейство таких связностей, получивших впоследствии название “связностей Ченцова–Амари”.

Н.Н. Ченцов рассмотрел единую модель ряда классических задач математической статистики как обратных задач теории вероятностей и нашел изящные и емкие формулировки асимптотических границ рисков в терминах размерности параметра, либо поперечников Колмогорова и внутренних радиусов Никольского априорных бесконечномерных семейств. Им установлена некорректность задачи оценки неизвестного распределения при отсутствии априорной информации.

Эти исследования подытожены в докторской диссертации “Общая теория статистического вывода”, защищенной в совете Института прикладной математики АН СССР в январе 1969 г.

В 70-е годы Н.Н. Ченцов совместно с Е.А. Морозовой начинает работать в новой, быстро развивающейся области — некоммутативной теории вероятностей, перенося свой категорно-геометрический подход на квантовую статистику. Была выяснена неполнота системы операций квантовой логики Биркгофа–Неймана и указана полная система операций и построена эргодическая теория для квантовых цепей Маркова. Одним из последних значительных достижений явилось описание целого класса инвариантных монотонных римановых метрик в пространстве квантовых состояний.

Тот факт, что уже в свои зрелые годы Н.Н. Ченцов сумел творчески освоить аппарат современной теоретической физики и плодотворно работать в такой новой и трудной для вероятностника, воспитанного в классических традициях теории множеств, области, как некоммутативная теория вероятностей, лишний раз свидетельствует о силе и свежести его математического дарования, столь безвременно утраченного.

Всю свою жизнь Н.Н. Ченцов оставался однолюбом и по отношению к выбранной научной стезе, и к семье, и к товарищам по работе, и к дорогому ему Институту прикладной математики им. М.В. Келдыша, где проработал без малого сорок лет. Он принимал самое деятельное участие в делах института — как ученый секретарь (что, по уставу АН, означало ближайшего помощника директора) и как заведующий отделом. Свои обязанности он исполнял весьма неформально.

5 июля 1992 г. после тяжелой болезни Николай Николаевич скончался.

Научная широта, глубокая эрудиция, энциклопедичность знаний, беззаветная преданность науке, необыкновенное трудолюбие, принципиальность, скромность и порядочность — вот те черты и качества, которые слились в образе Николая Николаевича Ченцова.

РАЗДЕЛ I

# ТЕОРИЯ СЛУЧАЙНЫХ ПРОЦЕССОВ И ПОЛЕЙ

---

## ВИНЕРОВСКИЕ СЛУЧАЙНЫЕ ПОЛЯ ОТ НЕСКОЛЬКИХ ПАРАМЕТРОВ \*)

*Представлено академиком А.Н. Колмогоровым 5 XI 1955*

В настоящей заметке мы дадим определение винеровского случайного поля от нескольких параметров — многомерного аналога винеровского случайного процесса, который, по нашей терминологии, является случайным полем от одного параметра. Это определение отлично от определения броуновского движения от  $n$  параметров, данного в [1]. Для краткости записи мы будем говорить только о полях от двух параметров  $t$  и  $s$ .

ОПРЕДЕЛЕНИЕ. *Винеровским случайным полем от двух параметров* будем называть сепарабельную случайную функцию двух действительных переменных  $W(t, s)$ , определенных на квадрате  $Q[0 \leq t, s \leq 1]$ , задаваемую многомерными нормальными распределениями с плотностью вероятности

$$\prod_{i=1}^M \prod_{j=1}^N [2\pi(t_i - t_{i-1})(s_j - s_{j-1})]^{-1/2} \times \\ \times \exp \left\{ - \frac{[x(t_i, s_j) - x(t_{i-1}, s_j) - x(t_i, s_{j-1}) + x(t_{i-1}, s_{j-1})]^2}{2(t_i - t_{i-1})(s_j - s_{j-1})} \right\}, \quad (1)$$

где  $0 = t_0 < t_1 < \dots < t_M = 1$ ;  $0 = s_0 < s_1 < \dots < s_N = 1$ ;

$$x(t_i, 0) \equiv 0; \quad x(0, s_j) \equiv 0,$$

или, в сокращенных обозначениях,

$$\prod_{i,j} [2\pi\Delta v_{ij}]^{-1/2} \exp \left\{ - \frac{(\Delta x_{ij})^2}{2\Delta v_{ij}} \right\}, \quad (1a)$$

---

\*) Докл. АН СССР. 1956. Т. 106. С. 4.607–609.

$\Delta x_{ij} = x(t_i, s_j) - x(t_{i-1}, s_j) - x(t_i, s_{j-1}) + x(t_{i-1}, s_{j-1})$  — смешанная разность и  $\Delta v_{ij} = (t_i - t_{i-1})(s_j - s_{j-1})$ .

Задаваемое таким образом случайное поле является случайным полем с независимыми приращениями по площадям (т.е. наше случайное поле является функцией распределения некоторой случайной меры на  $Q$  (ср. [2]) или решением дифференциального уравнения  $\partial^2 x / \partial t \partial s = \xi$ , где  $\xi(t, s)$  — обобщенная случайная нормально распределенная функция с независимыми значениями (ср. [3]). Поэтому совместные распределения, очевидно, согласованы между собой. Из определения характеристического функционала  $\varphi(\Lambda) = M[\exp\{i \int_Q W(t, s) d\Lambda\}]$  [4], где интеграл существует для всех  $W$ , если  $\Lambda(t, s)$  — функция распределения некоторой меры, сосредоточенной в конечном числе точек, легко выводится, что характеристический функционал винеровского случайного поля записывается в виде

$$\varphi(\Lambda) = \exp\left\{-\frac{1}{2} \int_0^1 \int_0^1 [\Lambda(t, s)]^2 dt ds\right\}, \quad (2)$$

или, сокращенно,

$$\varphi(\Lambda) = \exp\left\{-\frac{1}{2} \int_Q \Lambda^2 dv\right\},$$

где  $\Lambda(t, s)$  — ступенчатая функция и  $\Lambda(t, 1) = \Lambda(1, s) \equiv 0$ . (Заметим, что характеристический функционал винеровского процесса  $\exp\left\{-\frac{1}{2} \int_0^1 [\Lambda(t)]^2 dt\right\}$ , где  $\Lambda(t)$  — ступенчатая функция,  $\Lambda(1) = 0$ .)

Докажем следующий признак непрерывности случайного поля, обобщающий известный признак непрерывности процесса, принадлежащий Колмогорову, и выведем из него, что с вероятностью 1 винеровское случайное поле непрерывно.

**ТЕОРЕМА.** *Если случайное поле таково, что*

$$M|\Delta x|^\gamma < K|\Delta v|^{1+\varepsilon}, \quad \gamma > 0, \varepsilon > 0, \quad (3)$$

и

$$x(0, s) = x(t, 0) \equiv 0,$$

то с вероятностью 1 поле непрерывно.

**Доказательство.** Можно показать, что у процесса, обладающего свойством (3), множество сепарабельности может служить любое счетное всюду плотное множество. Мы докажем утверждение теоремы, предполагая, что множество сепарабельности состоит из всех двоично-рациональных точек квадрата  $Q$ . (Этот же метод доказательства применим и для случая любого множества сепарабельности  $\Gamma$ , надо только предварительно специальным образом отобразить множество  $t$ - и  $s$ -координат точек  $\Gamma$  на множество двоично-рациональных точек.)

Покажем, что с вероятностью, превышающей  $1 - \delta$ , имеет место событие  $A$  — все случайные поля равномерно ограничены величиной  $C(1 - \alpha)^{-2}$  и равномерно непрерывны с модулем непрерывности, не превышающим  $2C(1 - \alpha)^{-2}[\tau^{\varepsilon/2\gamma} + \delta^{\varepsilon/2\gamma}]$ , где  $C^\gamma = K/(\delta(1 - 2^{\varepsilon/2}))$ ,  $\alpha = 2^{-\varepsilon/2\gamma}$ .

Событие  $|\Delta x_{ij}^{mn}| < C\alpha^{n+m}$  будем называть событием  $B_{ij}^{mn}$ , где

$$\Delta x_{ij}^{mn} = x\left(\frac{i}{2^m}, \frac{j}{2^n}\right) - x\left(\frac{i-1}{2^m}, \frac{j}{2^n}\right) - x\left(\frac{i}{2^m}, \frac{j-1}{2^n}\right) + x\left(\frac{i-1}{2^m}, \frac{j-1}{2^n}\right).$$

Числа  $m$  и  $n$  мы будем называть *рангами отрезков*  $\left[\frac{i-1}{2^m}, \frac{i}{2^m}\right]$  и  $\left[\frac{j-1}{2^n}, \frac{j}{2^n}\right]$ . Чтобы оценить разность  $x(t'', s_0) - x(t', s_0)$ , надо просуммировать смешанные разности по какой-нибудь системе неперекрывающихся прямоугольников, заполняющих прямоугольник  $(t'', s_0); (t', s_0); (t', 0); (t'', 0)$ .

Пусть  $m_0$  — ранг наибольшего отрезка вида  $\left[\frac{i-1}{2^m}, \frac{i}{2^m}\right]$ , целиком лежащего в отрезке  $[t', t'']$ . Тогда отрезок  $[t', t'']$  можно представить как сумму не более двух отрезков ранга  $m_0$ , не более двух отрезков ранга  $m_0 + 1$  и т.д. Другую сторону можно представить как сумму отрезков различающихся между собой рангов. Разобьем весь прямоугольник на канонические прямоугольники согласно разбиению сторон. Пусть имеют место все события  $B_{ij}^{mn}$ . Тогда, оценивая модуль разности суммой модулей смешанных разностей, получаем, что равномерная ограниченность получается таким же путем. Оценивая точно так же  $|x(t, s'') - x(t, s')|$ , получаем

$$A \supset \bigcap_{m,n=0}^{\infty} \bigcap_{i,j} B_{ij}^{mn}.$$

Отсюда

$$P\{A\} > 1 - \sum_{i,j,m,n} [1 - P\{B_{ij}^{mn}\}].$$

Из условия теоремы выводится, что

$$1 - P\{B_{ij}^{mn}\} < KC^{-\gamma} \alpha^{-\gamma(m+n)} 2^{-(m+n)(1+\varepsilon)},$$

откуда, суммируя, получаем

$$P\{A\} > 1 - \delta.$$

Этим признак непрерывности доказан, так как  $\delta$  может быть взято сколь угодно малым.

Легко убедиться, что винеровское случайное поле удовлетворяет этому признаку, так как если  $\Delta x$  распределено нормально с параметрами  $(0, \sqrt{\Delta v})$ , то  $M|\Delta x|^4 = O(|\Delta v|^2)$ .

**ЗАМЕЧАНИЕ К ТЕОРЕМЕ.** Можно не требовать обращения поля в нуль на двух сторонах квадрата, а только предполагать, что случайные функции  $x(t, 0)$  и  $x(0, s)$  с вероятностью 1 непрерывны. Тогда в оценки войдут модули непрерывности этих функций.

Все сказанное выше тривиально переносится на случай, когда число  $p$  параметров больше двух, только в сокращенных формулах под  $\Delta x$  надо понимать  $p$ -ю смешанную разность, а под  $\Delta v$  — соответствующий объем. В доказательстве теоремы о непрерывности поля равномерно ограничены величиной  $C(1-\alpha)^{-p}$ , а модули непрерывности — функцией  $2C(1-\alpha)^{-p} \sum_{l=1}^p |\Delta t_l|^{\varepsilon/2\gamma}$ .

Аналогично могут рассматриваться векторные винеровские случайные поля от нескольких параметров.

## СПИСОК ЛИТЕРАТУРЫ

1. Levy P. // C. R. 1945. V. 220. P. 420.
2. Bocher S. // Proc. Nat. Acad. Sci. USA. 1946. V. 32. P. 259.
3. Гельфанд И.М. // Докл. АН СССР. 1955. Т. 100, № 5. С. 853.
4. Колмогоров А.Н. // С.Р. 1935. V. 200. P. 1717.

**СЛАБАЯ СХОДИМОСТЬ СЛУЧАЙНЫХ  
ПРОЦЕССОВ С ТРАЕКТОРИЯМИ  
БЕЗ РАЗРЫВОВ ВТОРОГО РОДА  
И ТАК НАЗЫВАЕМЫЙ “ЭВРИСТИЧЕСКИЙ”  
ПОДХОД К КРИТЕРИЯМ СОГЛАСИЯ  
ТИПА КОЛМОГОРОВА–СМИРНОВА <sup>\*</sup>)**

В 1934 г. А.Н. Колмогоров доказал следующую теорему. Если  $\xi(t)$  — сепарабельный (см. [1]) случайный процесс,

$$\mathbb{M} |\xi(t_1) - \xi(t_2)|^p < C |t_1 - t_2|^{1+r}, \quad (1)$$

где  $p > 0$ ,  $r > 0$ ,  $C$  — константа, не зависящая от  $t$ , то с вероятностью 1 все траектории процесса непрерывны<sup>1)</sup>. Обобщением этой теоремы является следующее предложение, указанное автору А.Н. Колмогоровым.

ТЕОРЕМА 1. Если  $\xi(t)$  — сепарабельный случайный процесс  $0 \leq t \leq 1$  и

$$\mathbb{M} |\xi(t_1) - \xi(t_2)|^p |\xi(t_2) - \xi(t_3)|^q < C |t_1 - t_3|^{1+r}, \quad (2)$$

где  $p \geq 0$ ,  $q \geq 0$ ;  $r > 0$ ;  $t_1 < t_2 < t_3$ ,  $C$  — константа, не зависящая от  $t$ , то с вероятностью 1 все траектории процесса не имеют разрывов второго рода.

Доказательство. Нетрудно показать, что если для  $\xi(t)$  выполнено (2), в качестве множества сепарабельности можно выбрать любое счетное всюду плотное множество, например множество  $\Gamma_0$  двоично-рациональных дробей. Назовем событие

$$\left| \xi\left(\frac{k}{2^n}\right) - \xi\left(\frac{k+1}{2^n}\right) \right| < L\alpha^n,$$

где  $\alpha = \exp\left[-\frac{r \ln 2}{2(p+q)}\right]$ ,  $L^{p+q} = C$ , событием  $A_{k+1/2}^n$ . Введем также события

$$B_k^n = A_{k-1/2}^n \cup A_{k+1/2}^n, \text{ события } D^n = \bigcap_{m=n}^{\infty} \bigcap_{k=1}^{2^m-1} B_k^m \text{ и событие } D = \bigcup_{n=1}^{\infty} D^n.$$

<sup>\*</sup>) Теория вероятн. и ее примен. 1956. Т. 1, вып. 1. С. 155–161.

<sup>1)</sup> Эта теорема была впервые опубликована в работе Е.Е. Слуцкого [2].

Из неравенства (2) легко выводится, что  $\mathbf{P}\{\overline{B}_k^n\} < \alpha^{-n(p+q)}2^{-n(1+r)}$ ,  $\mathbf{P}\{\overline{D}^n\} < \varepsilon^n/(1-\varepsilon)$ , где  $\varepsilon = 2^{-r/2} < 1$ , и, наконец,  $\mathbf{P}\{D\} = 1$ . Остается показать, что множество  $D$  не содержит ни одной  $\Gamma_0$ -сепарабельной траектории с разрывами второго рода.

ЛЕММА 1. Если  $x(t)$  —  $\Gamma_0$ -сепарабельная функция и  $x(t) \in A_{k+1/2}^n \cap D^{n+1}$ , то

$$\max_{\frac{k}{2^n} < t < \frac{k+1}{2^n}} \left| x(t) - x\left(\frac{k}{2^n}\right) \right| < \frac{L\alpha^n}{1-\alpha}, \quad (3)$$

$$\max_{\frac{k}{2^n} < t < \frac{k+1}{2^n}} \left| x(t) - x\left(\frac{k+1}{2^n}\right) \right| < \frac{L\alpha^n}{1-\alpha}. \quad (4)$$

Доказательство. Так как  $x(t)$  —  $\Gamma_0$ -сепарабельная функция, то утверждение леммы достаточно показать для двоично-рациональных  $t$ . Пусть по предположению индукции

$$\left| x\left(\frac{S}{2^{n+m}}\right) - x\left(\frac{k}{2^n}\right) \right| < L(\alpha^n + \alpha^{n+1} + \dots + \alpha^{n+m}), \quad (5)$$

$$\left| x\left(\frac{S}{2^{n+m}}\right) - x\left(\frac{k+1}{2^n}\right) \right| < L(\alpha^n + \alpha^{n+1} + \dots + \alpha^{n+m}). \quad (6)$$

Так как  $x(t) \in D^{n+1}$ , то имеет место по крайней мере одно из неравенств

$$\left| x\left(\frac{2S+1}{2^{n+m+1}}\right) - x\left(\frac{S}{2^{n+m}}\right) \right| < L\alpha^{n+m+1}, \quad (7)$$

$$\left| x\left(\frac{2S-1}{2^{n+m+1}}\right) - x\left(\frac{S}{2^{n+m}}\right) \right| < L\alpha^{n+m+1}. \quad (8)$$

В обоих случаях

$$\left| x\left(\frac{2S+1}{2^{n+m+1}}\right) - x\left(\frac{k}{2^n}\right) \right| < L(\alpha^n + \alpha^{n+1} + \dots + \alpha^{n+m+1}), \quad (9)$$

$$\left| x\left(\frac{2S+1}{2^{n+m+1}}\right) - x\left(\frac{k+1}{2^n}\right) \right| < L(\alpha^n + \alpha^{n+1} + \dots + \alpha^{n+m+1}). \quad (10)$$

Дальнейшие рассуждения тривиальны.

ЛЕММА 2. Если  $x(t)$  —  $\Gamma_0$ -сепарабельная функция и  $x(t) \in B_k^n \cap D^{n+1}$ , то в интервале  $[(k-1)/2^n, (k+1)/2^n]$  существует точка  $t_k^n$ , задающая дедкиндово сечение интервала так, что

$$\left| x(t) - x\left(\frac{k-1}{2^n}\right) \right| < \frac{L\alpha^n}{1-\alpha}, \quad (11)$$

если  $t$  принадлежит нижнему классу, и

$$\left| x(t) - x\left(\frac{k+1}{2^n}\right) \right| < \frac{L\alpha^n}{1-\alpha}, \quad (12)$$

если  $t$  принадлежит верхнему классу.

Доказательство. Всегда имеет место одно из неравенств

$$\left| x\left(\frac{k}{2^n}\right) - x\left(\frac{k-1}{2^n}\right) \right| < L\alpha^n, \quad (13)$$

$$\left| x\left(\frac{k}{2^n}\right) - x\left(\frac{k+1}{2^n}\right) \right| < L\alpha^n. \quad (14)$$

В случае когда выполнены оба неравенства, утверждение леммы 2 тривиально следует из леммы 1. Пусть выполнено только одно неравенство, например первое. Тогда из леммы 1 следует, что для всех  $t \in [(k-1)/2^n, k/2^n]$  выполнено неравенство (11), поэтому отрезок  $[(k-1)/2^n, k/2^n]$  принадлежит нижнему классу. Для концов отрезка  $[k/2^n, (k+1)/2^n]$  имеем

$$\left| x\left(\frac{k}{2^n}\right) - x\left(\frac{k-1}{2^n}\right) \right| < L\alpha^n, \quad (15)$$

$$\left| x\left(\frac{k+1}{2^n}\right) - x\left(\frac{k+1}{2^n}\right) \right| = 0 < L\alpha^n. \quad (16)$$

Отрезок  $[k/2^n, (k+1)/2^n]$  разбиваем пополам. Так как имеет место событие  $B_{k+1}^{n+1}$ , то хотя бы для одной из половинок имеем

$$\left| x\left(\frac{S}{2^{n+1}}\right) - x\left(\frac{S+1}{2^{n+1}}\right) \right| < L\alpha^{n+1},$$

где  $S = 2k$  или  $S = 2k + 1$ . Из леммы 1 и одного из неравенств (15) или (16) получаем, что все эти точки  $[S/2^{n+1}, (S+1)/2^{n+1}]$  принадлежат одному классу. Для второй половины отрезка получаем оценки

$$\left| x\left(\frac{S'}{2^{n+1}}\right) - x\left(\frac{k-1}{2^n}\right) \right| < L(\alpha^n + \alpha^{n+1}), \quad \left| x\left(\frac{k+1}{2^{n+1}}\right) - x\left(\frac{k+1}{2^n}\right) \right| < L(\alpha^n + \alpha^{n+1}),$$

где  $S' = 2k + 1$ , если  $S = 2k$ , или  $S' = 2k$ , если  $S = 2k + 1$ . Рассуждая далее по индукции, получаем утверждение леммы.

Переходим к доказательству теоремы. Пусть  $x(t)$  — функция с разрывами второго рода и  $x(t) \in D$ . Тогда при некотором  $n$ ,  $x(t) \in D^n$  и, следовательно,  $x(t) \in D^m$  при всех  $m \geq n$ . Если  $\Gamma_0$ -сепарабельная функция  $x(t)$  имеет разрывы второго рода, то существует такое  $\eta > 0$ , что для любого  $\delta > 0$  найдутся три точки  $t_1 < t_2 < t_3$   $|t_1 - t_3| < \delta$ ,  $t_1, t_2, t_3 \in \Gamma_0$  такие, что  $|x(t_1) - x(t_2)| > \eta$  и  $|x(t_2) - x(t_3)| > \eta$ .

Выберем  $m$  так, чтобы  $2L\alpha^m/(1-\alpha) < \eta$ . Без ограничения общности можно считать  $m \geq n$ . Выберем  $\delta < 1/2^m$ . Тогда отрезок  $[t_1, t_3]$  заведомо покрывается двумя смежными отрезками длины  $[(k-1)/2^m, k/2^m]/2^m$  и  $[k/2^m, (k+1)/2^m]$ . При этом или  $t_1$  и  $t_2$ , или  $t_2$  и  $t_3$ , или и  $t_1$  и  $t_2$  и  $t_3$  попадают в один класс. Так как  $x(t) \in B_k^m \cap D^{m+1}$ , то по следствию из леммы 2 одна из разностей обязательно меньше  $2L\alpha^m/(1-\alpha)$ , что противоречит неравенству  $2L\alpha^m/(1-\alpha) < \eta$ . Это противоречие доказывает теорему 1.

Будем говорить, что последовательность случайных процессов  $\xi_n(t)$  слабо сходится к некоторому предельному случайному процессу  $\xi_0(t)$ , если при любом фиксированном множестве моментов времени  $t_1, t_2, \dots, t_m$  последовательность совместных распределений величин  $\xi_n(t_1), \dots, \xi_n(t_m)$  слабо сходится к соответствующему совместному распределению значений  $\xi_0(t_i), i = 1, \dots, m$ , предельного процесса.

ТЕОРЕМА 2. Если последовательность сепарабельных случайных процессов  $\xi_n(t)$  слабо сходится к сепарабельному случайному процессу  $\xi_0(t)$ , причем выполнено условие

$$\mathbf{M} |\xi_n(t_1) - \xi_n(t_2)|^p |\xi_n(t_2) - \xi_n(t_3)|^q < C |t_1 - t_3|^{1+r}, \quad (17)$$

где  $p \geq 0, q \geq 0; r \geq 0; t_1 < t_2 < t_3; C$  — константа, не зависящая от  $n$  и  $t$ , и две кусочно непрерывные функции  $g(t) < f(t)$ , полунепрерывные в точках разрыва соответственно сверху и снизу, таковы, что

$$\mathbf{P}\{g(t)+z \leq \xi_0(t) \leq f(t)-z \text{ при всех } t\} \xrightarrow{z \rightarrow +0} \mathbf{P}\{g(t) \leq \xi_0(t) \leq f(t) \text{ при всех } t\}, \quad (18)$$

то

$$\mathbf{P}\{g(t) \leq \xi_n(t) \leq f(t) \text{ при всех } t\} \xrightarrow{n \rightarrow \infty} \mathbf{P}\{g(t) \leq \xi_0(t) \leq f(t) \text{ при всех } t\}. \quad (19)$$

Доказательство. Если процесс  $\xi(t)$  сепарабельный, то (см. [1])

$$\mathbf{P}\{g(t) \leq \xi(t) \leq f(t) \text{ при всех } t \in T\} = \inf_{\Gamma} \mathbf{P}\{g(t) \leq \xi(t) \leq f(t) \text{ при всех } t \in \Gamma\}, \quad (20)$$

где  $\Gamma$  — любое счетное подмножество  $T$  и  $g(t) < f(t)$  — любые кусочно непрерывные функции. Под  $T$  для определенности будем понимать отрезок оси  $t$ . Пусть счетное всюду плотное множество  $\Gamma_0$  содержит все точки разрыва функций  $d(t)$  и  $f(t)$ . Нетрудно показать, что если для  $\xi(t)$  выполнено условие (17), то на  $\Gamma_0$  достигается  $\inf_{\Gamma} \mathbf{P}\{g(t) \leq \xi(t) \leq f(t) \text{ при всех } t \in \Gamma\}$ . В самом деле, из (17) следует, что от увеличения  $\Gamma_0$  на любое конечное, а значит, и на счетное число точек эта вероятность не будет уменьшаться. Так как, очевидно, условие (17) имеет место и для предельного процесса  $\xi_0(t)$ , то нам достаточно показать, что

$$\mathbf{P}\{g(t) \leq \xi_n(t) \leq f(t) \text{ при всех } t \in \Gamma_0\} \xrightarrow{n \rightarrow \infty} \mathbf{P}\{g(t) \leq \xi_0(t) \leq f(t) \text{ при всех } t \in \Gamma_0\}. \quad (21)$$

Выберем  $\Gamma_0$  следующим специальным образом: пусть  $\Delta_0$  состоит из точек разрыва функции  $g(t)$  и  $f(t)$  и концов отрезка  $T$  оси  $t$ . Пусть  $\Delta_1$  состоит из всех точек  $\Delta_0$  и всех точек вида  $(t_{i+1}^{(0)} + t_i^{(0)})/2$ , где  $t_i^{(0)}$  — упорядоченные в порядке возрастания точки  $\Delta_0$ . Пусть  $\Delta_2$  состоит из всех точек  $\Delta_1$  и всех точек вида  $(t_{i+1}^{(1)} + t_i^{(1)})/2$ , где  $t_i^{(1)}$  — упорядоченные в порядке возрастания точки  $\Delta_1$  и т.д. Положим, наконец,  $\Gamma_0 = \bigcup_{k=1}^{\infty} \Delta_k$ .

Обозначим для краткости

$$\{g(t) \leq \xi(t) \leq f(t) \text{ при всех } t \in A\} = \mathbf{U}(\xi, g, f, A). \quad (22)$$

Тогда (21) можно переписать в виде

$$\mathbf{P}\{\mathbf{U}(\xi_n, g, f, \Gamma_0)\} \xrightarrow{n \rightarrow \infty} \mathbf{P}\{\mathbf{U}(\xi_0, g, f, \Gamma_0)\}. \quad (23)$$

Пусть  $g(t)$  и  $f(t)$  — ступенчатые функции. Тогда имеет место

ЛЕММА. Если для  $\xi(t)$  выполнено условие (17), то при  $\varepsilon \geq \alpha^k / (1 - \alpha)$

$$\mathbf{P}\{\mathbf{U}(\xi, g, f, \Delta_k)\} \leq \mathbf{P}\{\mathbf{U}(\xi, g - \varepsilon, f + \varepsilon, \Gamma_0)\} + \eta, \quad (24)$$

где

$$h = \max_i |t_{i+1}^{(0)} + t_i^{(0)}|, \quad \alpha = \exp \left[ -\frac{r \ln 2}{2(p+q)} \right] < 1, \quad \eta \leq C |T| h^r \frac{\alpha^{(p+q)k}}{1 - \alpha^{p+q}}.$$

Доказательство. Нетрудно проверить, что из (17) следует

$$\mathbf{P}\left\{ \xi(t_i^{(l)}) \in [a, b]; \xi(t_{i+1}^{(l)}) \in [a, b]; \xi(t_{2i}^{(l+1)}) \in [a - \varepsilon, b + \varepsilon] \right\} \leq \frac{C |t_{i+1}^{(l)} + t_i^{(l)}|^{1+r}}{\varepsilon^{p+q}}. \quad (25)$$

В силу полунепрерывности  $g(t)$  и  $f(t)$  в точках разрыва можно подставить в (25) вместо  $a$  величину  $g(t_{2i}^{(l+1)}) - \varepsilon_l$  и  $g(t_{2i}^{(l+1)}) + \varepsilon_l$  вместо  $b$ , где  $\varepsilon_l = \alpha^k + \alpha^{k+1} + \dots + \alpha^{l+1}$ . Суммируя затем по всем  $i$ , получаем

$$\mathbf{P}\{\mathbf{U}(\xi, g - \varepsilon_l, f + \varepsilon_l, \Delta_l)\} \leq \mathbf{P}\{\mathbf{U}(\xi, g - \varepsilon_{l+1}, f + \varepsilon_{l+1}, \Delta_{l+1})\} + \eta_l, \quad (26)$$

где  $\eta_l \leq C h^r |T| \alpha^{k(p+q)l}$ . Суммируя выражения (26) для всех  $l \geq k$  и учитывая, что от увеличения  $\varepsilon$  неравенство только усиливается, получаем (24). Из леммы следует, что для любых  $\varepsilon > 0$  и  $\eta > 0$  найдется такое  $k$ , зависящее только от  $C$ ,  $h$ ,  $|T|$ ,  $p$ ,  $q$ ,  $r$ , что неравенство (24) будет выполнено с этими  $\varepsilon$  и  $\eta$ .

Вернемся к доказательству теоремы. Выберем произвольное  $\eta > 0$  и покажем, что

$$\overline{\lim}_{n \rightarrow \infty} \mathbf{P}\{\mathbf{U}(\xi_n, g, f, \Gamma_0)\} < \mathbf{P}\{\mathbf{U}(\xi_0, g, f, \Gamma_0)\} + \eta. \quad (27)$$

Так как  $\mathbf{U}(\xi, g, f, \Gamma_0) = \bigcap_{k=1}^{\infty} \mathbf{U}(\xi, g, f, \Delta_k)$ , то можно выбрать  $k$  таким, что

$$\mathbf{P}\{\mathbf{U}(\xi_0, g, f, \Delta_k)\} < \mathbf{P}\{\mathbf{U}(\xi_0, g, f, \Gamma_0)\} + \frac{\eta}{2}. \quad (28)$$

Выберем теперь  $n_0$  таким, что при  $n \geq n_0$

$$\mathbf{P}\{\mathbf{U}(\xi_n, g, f, \Delta_k)\} < \mathbf{P}\{\mathbf{U}(\xi_0, g, f, \Delta_k)\} + \frac{\eta}{2}, \quad (29)$$

что можно сделать, так как множество  $\mathbf{U}(\xi, g, f, \Delta_k)$  замкнуто. Наконец, всегда имеет место

$$\mathbf{P}\{\mathbf{U}(\xi_n, g, f, \Gamma_0)\} \leq \mathbf{P}\{\mathbf{U}(\xi_n, g, f, \Delta_k)\}. \quad (30)$$

Складывая (28), (29) и (30) и переходя к пределу по  $n$ , получаем (27).

Заметим, что при доказательстве неравенства (27) мы не накладывали никаких ограничений на функции  $g(t)$  и  $f(t)$ .

Сложнее доказательство обратного неравенства

$$\lim_{n \rightarrow \infty} \mathbf{P}\{\mathbf{U}(\xi_n, g, f, \Gamma_0)\} > \mathbf{P}\{\mathbf{U}(\xi_0, g, f, \Gamma_0)\} - \eta. \quad (31)$$

Согласно (18) всегда можно выбрать  $\varepsilon > 0$  так, что

$$\mathbf{P}\{\mathbf{U}(\xi_0, g + 3\varepsilon, f - 3\varepsilon, \Gamma_0)\} > \mathbf{P}\{\mathbf{U}(\xi_0, g, f, \Gamma_0)\} - \frac{\eta}{2}. \quad (32)$$

Каковы бы ни были кусочно непрерывные функции  $g(t)$  и  $f(t)$ , всегда можно выбрать ступенчатые функции  $g_1(t)$  и  $f_1(t)$  с разрывами в точках  $\Gamma_0$  так, чтобы  $g_1(t) + \varepsilon > g(t) > g_1(t)$ ;  $f_1(t) > f(t) > f_1(t) - \varepsilon$ . Тогда

$$\mathbf{U}(\xi, g_1 + \delta + \varepsilon, f_1 - \delta - \varepsilon, \Gamma_0) \subset \mathbf{U}(\xi, g + \delta, f - \delta, \Gamma_0) \subset \mathbf{U}(\xi, g_1 + \delta, f_1 - \delta, \Gamma_0). \quad (33)$$

Следовательно, имеют место неравенства

$$\mathbf{P}\{\mathbf{U}(\xi_0, g_1 + 3\varepsilon, f_1 - 3\varepsilon, \Gamma_0)\} \geq \mathbf{P}\{\mathbf{U}(\xi_0, g + 3\varepsilon, f - 3\varepsilon, \Gamma_0)\}, \quad (34)$$

$$\mathbf{P}\{\mathbf{U}(\xi_n, g, f, \Gamma_0)\} \geq \mathbf{P}\{\mathbf{U}(\xi_n, g_1 + \varepsilon, f_1 - \varepsilon, \Gamma_0)\}. \quad (35)$$

Выберем по  $\varepsilon$  такое  $k$ , чтобы при любом  $n$  имело место

$$\mathbf{P}\{\mathbf{U}(\xi_n, g_1 + \varepsilon, f_1 - \varepsilon, \Gamma_0)\} > \mathbf{P}\{\mathbf{U}(\xi_n, g_1 + 2\varepsilon, f_1 - 2\varepsilon, \Delta_k)\} - \frac{\eta}{2}. \quad (36)$$

Возможность такого выбора доказана в лемме. Выберем теперь  $n_0$  так, что при  $n \geq n_0$

$$\mathbf{P}\{\mathbf{U}(\xi_n, g_1 + 2\varepsilon, f_1 - 2\varepsilon, \Delta_k)\} > \mathbf{P}\{\mathbf{U}(\xi_0, g_1 + 3\varepsilon, f_1 - 3\varepsilon, \Delta_k)\}. \quad (37)$$

Наконец, всегда имеет место

$$\mathbf{P}\{\mathbf{U}(\xi_0, g_1 + 3\varepsilon, f_1 - 3\varepsilon, \Delta_k)\} \geq \mathbf{P}\{\mathbf{U}(\xi_0, g_1 + 3\varepsilon, f_1 - 3\varepsilon, \Gamma_0)\}. \quad (38)$$

Складывая соотношения (32), (34)–(38) и переходя по  $n$  к пределу, получаем (31). Так как  $\eta$  произвольно, то  $\lim_{n \rightarrow \infty} \mathbf{P}\{\mathbf{U}(\xi_n, g, f, \Gamma_0)\}$  существует и равен  $\mathbf{P}\{\mathbf{U}(\xi_0, g, f, \Gamma_0)\}$ .

ЗАМЕЧАНИЯ. 1. В доказательстве фактически используется не условие (17), а менее ограничительное соотношение (25).

2. Доказательство легко переносится на случай, когда  $g(t)$  и  $f(t)$  суть равномерные монотонные пределы ступенчатых функций, полунепрерывных соответственно сверху и снизу, т.е. функции без разрывов второго рода, полунепрерывные соответственно сверху и снизу.

3. Аналогичную теорему можно доказать для векторных случайных процессов, а также для случайных полей.

4. Пусть известно, что процесс  $\xi(t)$  стохастически непрерывен в точке  $t_0$  и удовлетворяет условию (17). Положим в формуле (24)  $k = 0$ ,  $|T| = h$  и заменим  $\xi(t)$  на случайную функцию  $\xi(t) - \xi(t_0)$ , также удовлетворяющую условию (17). Из этой оценки нетрудно вывести, что процесс  $\xi(t)$  непрерывен в точке  $t_0$  с вероятностью 1.

Доказанная теорема 2 дает новое обоснование метода разыскания асимптотических распределений критериев согласия, предложенного в нестрогой, эвристической форме Дубом [3] и развитого затем Донскером [4]. (Метод этот был близок к тому, который был использован Колмогоровым [5] еще в 1933 г.) Рассмотрим случайную функцию

$$\gamma_n(t) = \sqrt{n} [F_n^*(t) - t],$$

где  $F_n^*(t)$  — эмпирическая функция распределения  $n$  наблюдений случайной величины, равномерно распределенной на отрезке  $[0, 1]$ . Если мы введем вспомогательную функцию распределения  $\Lambda(t)$  меры, сосредоточенной в точках  $t_1, t_2, \dots, t_m$  с массами соответственно  $-\lambda_1, -\lambda_2, \dots, -\lambda_m$ , причем будем считать, что  $\Lambda(1) = 0$ , то характеристическую функцию совместного распределения значений случайного “процесса”  $\gamma_n(t)$  в “моменты времени”  $t_1, \dots, t_m$  удобно записать в виде

$$\varphi_n(\lambda_1, \lambda_2, \dots, \lambda_m) = \left[ \exp \left\{ -i \int_0^1 \frac{1}{\sqrt{n}} \Lambda(t) dt \right\} \int_0^1 \exp \left\{ \frac{i}{\sqrt{n}} \Lambda(t) \right\} dt \right]^n. \quad (39)$$

Нетрудно видеть, что при  $n \rightarrow \infty$  и при фиксированных  $t_1, \dots, t_m$  и  $\lambda_1, \dots, \lambda_m$  функции  $\varphi_n(\lambda_1, \dots, \lambda_m)$  сходятся к характеристической функции условного винеровского процесса  $\xi(t)$  с условием  $\xi(1) = 0$ , которая может быть записана в виде

$$\varphi(\lambda_1, \lambda_2, \dots, \lambda_m) = \exp \left\{ -\frac{1}{2} \int_0^1 [\Lambda(t)]^2 dt + \frac{1}{2} \left[ \int_0^1 \Lambda(t) dt \right]^2 \right\}. \quad (40)$$

Нетрудно проверить, что процесс  $\gamma_n(t)$  удовлетворяет условию (17), так как

$$\begin{aligned} \mathbf{M} [\gamma_n(t + \tau_1) - \gamma_n(t)]^2 [\gamma_n(t) - \gamma_n(t - \tau_2)]^2 &= \left(1 - \frac{1}{n}\right) \tau_1 \tau_2 (1 - \tau_1)(1 - \tau_2) + \\ &+ \frac{1}{n} \tau_2 \tau_2 [\tau_2 (1 - \tau_1)^2 + \tau_1 (1 - \tau_2)^2 + \tau_1 \tau_2 (1 - \tau_1 - \tau_2)] = O(|\tau_1 + \tau_2|^2). \end{aligned}$$

Поэтому мы можем утверждать, что асимптотическое распределение любого критерия согласия типа критериев согласия Колмогорова [5] и Смирнова [6] (см., например, [7, 8]) совпадает с соответствующим распределением для условного винеровского процесса.

Следует отметить, что точно так же можно свести доказательство предложений о поведении последовательности сумм независимых случайных величин к доказательству аналогичных предложений для винеровского случайного процесса.

В заключение автор выражает глубокую благодарность Андрею Николаевичу Колмогорову и Николаю Васильевичу Смирнову, руководившим его работой.

## СПИСОК ЛИТЕРАТУРЫ

1. Doob J.L. Stochastic processes. — N.Y.—London, 1953.
2. Слуцкий Е.Е. // Giornale dell'Istituto Italiano degli Attuari. 1937. V. 8. P. 183–199.
3. Doob J.L. // Ann. Math. Statistics. 1949. V. 20. P. 393–403.
4. Donsker M.D. // Ann. Math. Statistics. 1952. V. 23. P. 277–281.
5. Колмогоров А.Н. // Giornale dell'Istituto Italiano degli Attuari. 1933. V. 4. P. 83–91.
6. Смирнов Н.В. // Матем. сб. 1939. Т. 6 (44). С. 3–26.
7. Манья Г.М. // Докл. АН СССР. 1949. Т. 69. С. 495–497.
8. Реньи А. // Acta Math. Acad. Sci. Hungaricae. 1953. V. 4. P. 191–231.

## МНОГОПАРАМЕТРИЧЕСКОЕ БРОУНОВСКОЕ ДВИЖЕНИЕ ЛЕВИ И ОБОБЩЕННЫЙ БЕЛЫЙ ШУМ<sup>\*)</sup>

В настоящей заметке дается наглядное геометрическое истолкование многопараметрического броуновского движения Леви [1]. Для большей элементарности изложения мы будем проводить рассуждения, не опираясь на общие теоремы работы Ито [2].

Пусть задано однородное пространство с инвариантной  $\sigma$ -конечной мерой  $\mu$ . Следуя Бохнеру [3], будем называть *белым шумом*  $\xi(V)$  случайную функцию, заданную на множествах конечной меры из некоторой  $\sigma$ -алгебры и обладающую следующими свойствами:

- а)  $\xi(V)$  распределено нормально со средним 0 и дисперсией  $\mu(V)$ ;
- б) если множества  $V$  и  $W$  не пересекаются, то  $\xi(V)$  и  $\xi(W)$  независимы и  $\xi(V \cup W) = \xi(V) * \xi(W)$ .

Можно сказать, что белый шум — это случайная мера, принимающая значения разных знаков и имеющая, вообще говоря, неограниченную вариацию<sup>1)</sup>.

Будем рассматривать евклидову плоскость как совокупность прямых. Как известно, в этом пространстве существует мера, инвариантная относительно группы движений. (Если условиться задавать прямую в виде  $x \cos \varphi + y \sin \varphi - p = 0$ , где  $0 \leq \varphi \leq 2\pi$ ;  $p \geq 0$ , то эта мера множества  $V$  прямых задается интегралом  $\mu(V) = \int_V d\varphi dp$ ; см., например, [6].) Вычислим меру некоторых множеств, не используя ее аналитического задания. Очевидно, в силу  $\sigma$ -конечности и инвариантности мера одной прямой равна нулю; точно по тем же соображениям равна нулю мера пучка прямых, проходящих через фиксированную точку, или мера пучка параллельных прямых. Так как меры множеств прямых, пересекающих равные отрезки, должны быть одинаковы, то мера количества прямых,

<sup>\*)</sup> Теория вероятн. и ее примен. 1957. Т. 2, вып. 2. С. 281–282.

<sup>1)</sup> Определение белого шума было дано Гельфандом [4] и Ито [5] в несколько более общей форме.

пересекающих данный отрезок, пропорциональна его длине. Нам удобно выбрать коэффициент пропорциональности равным единице.

Обозначим  $\widehat{A}$ ,  $\widehat{AB}$  соответственно множества прямых, проходящих через точку  $A$  и пересекающих отрезок  $AB$ . Из геометрических аксиом имеем

$$\begin{aligned} & (\widehat{AB} \cap \widehat{BC}) \cup (\widehat{AB} \cap \widehat{CA}) = \widehat{AB}, \\ & (\widehat{AB} \cap \widehat{BC}) \cap (\widehat{AB} \cap \widehat{CA}) \subset \widehat{A} \cup \widehat{B} \cup \widehat{C}. \end{aligned}$$

Поэтому

$$\mu(\widehat{AB} \cap \widehat{BC}) + \mu(\widehat{AB} \cap \widehat{CA}) = \mu(\widehat{AB}).$$

Рассмотрим белый шум в пространстве прямых евклидовой плоскости. Выберем какую-нибудь точку  $O$  за начальную и положим для каждой точки  $A$ , что  $\eta(A) = \xi(\widehat{OA})$ .

Леви определял броуновское движение от  $m$  параметров как случайное поле  $X(A)$ , обладающее следующими свойствами:

- а) величины  $X(A_j) = X(A_j) - X(0)$  при любом выборе точек  $A_1, \dots, A_n$ ,  $n = 1, 2, \dots$ , имеют совместное несмещенное нормальное распределение;
- б)  $M |X(A) - X(B)|^2 = |AB|$ .

Эти условия определяют случайное поле однозначно, с точностью до выбора точки  $O$ . Очевидно, что наше случайное поле  $\eta(A)$  обладает свойством а), так как мы можем разбить множества  $\widehat{OA}_i$ ,  $i = 1, \dots, n$ , на  $2^n$  не пересекающихся между собой множеств и соответственно этому представить  $\eta(A_i) = \xi(\widehat{OA}_i)$ ,  $i = 1, \dots, n$ , как сумму независимых нормальных случайных величин с математическим ожиданием, равным нулю. Подобным же приемом можно вычислить дисперсию:

$$\begin{aligned} \eta(A) - \eta(B) &= \xi(\widehat{OA}) - \xi(\widehat{OB}) = \\ &= \xi(\widehat{OA} \cap \widehat{AB}) + \xi(\widehat{OA} \cap \widehat{OB}) - \xi(\widehat{OB} \cap \widehat{AB}) - \xi(\widehat{OB} \cap \widehat{OA}) = \\ &= \xi(\widehat{OA} \cap \widehat{AB}) - \xi(\widehat{OB} \cap \widehat{AB}). \end{aligned}$$

Так как  $\mu(\widehat{OA} \cap \widehat{AB} \cap \widehat{OB}) = 0$ , то  $\xi(\widehat{OA} \cap \widehat{AB})$  и  $\xi(\widehat{OB} \cap \widehat{AB})$  независимы и дисперсия их разности равна сумме дисперсий, т.е.

$$\mathbf{D}[\eta(A) - \eta(B)] = \mu(\widehat{OA} \cap \widehat{AB}) + \mu(\widehat{OB} \cap \widehat{AB}) = |AB|.$$

Таким образом, построенное наше случайное поле  $\eta(A)$  является броуновским движением Леви от двух параметров. Все рассуждения и выкладки дословно переносятся и на случай большего числа измерений, только в  $m$ -мерном пространстве надо рассматривать инвариантную меру множества  $(m-1)$ -мерных гиперплоскостей, пересекающих данный отрезок.

## СПИСОК ЛИТЕРАТУРЫ

1. *Levy P.* Processus stochastiques et mouvement brownien. — Paris: Gauthier-Villars, 1948.
2. *Itô K.* Isotropic random current // Proc. 3-rd Berkeley Symp. on Math. Statistics and Probability. — University of California Press.
3. *Bochner S.* Harmonic analysis and the theory of probability. — 1955.
4. *Гельфанд И.М.* // Докл. АН СССР. 1955. Т. 100. С. 853–856.
5. *Itô K.* Complex multiple Wiener integral // Jap. J. Math. 1952. V. 22. P. 63–86.
6. *Santaló L.A.* Introduction to integral geometry. — Paris: Hermann & Cie, 1953. [Рус. пер.: *Сантало Л.А.* Введение в интегральную геометрию. — М.: ИЛ, 1956.]

## ТОПОЛОГИЧЕСКИЕ МЕРЫ И ТЕОРИЯ СЛУЧАЙНЫХ ФУНКЦИЙ<sup>\*)</sup>

Пусть монотонная функция множества  $\mu(\cdot)$  задана на алгебре  $S$  подмножеств топологического пространства  $U$ . Обозначим через  $S_g$  и  $S_f$  соответственно структуры открытых и замкнутых множеств, входящих в  $S$ .

ОПРЕДЕЛЕНИЕ 1. Будем говорить, что функция  $\mu(\cdot)$  является *топологической*, если для любого  $A \in S$

$$\mu(A) = \inf_{S_g \ni G \supseteq A} \mu(G) = \sup_{S_f \ni F \subseteq A} \mu(F). \quad (1)$$

Если пространство  $U$  имеет несчетную базу, то в  $U$  приходится отличать бэровские и борелевские множества; см., например, [1]. Класс всех борелевских множеств, как обычно, является  $\sigma$ -алгеброй  $B$ -множеств, порожденных классом всех открытых подмножеств  $U$ . Класс  $B_0$  бэровских множеств является  $\sigma$ -алгеброй, порожденной открытыми множествами типа  $F_\sigma$ . Всякая бэровская мера, как нетрудно проверить, автоматически является топологической.

Пусть  $\xi(t)$  — случайная функция аргумента  $t \in T$  со значениями в топологическом пространстве  $X$ . Тогда  $\xi(\cdot)$  индуцирует распределение вероятностей (т.е. меру) в топологическом пространстве  $X^T$  — тихоновской топологической степени пространства  $X$ . В том случае, когда  $X$  — локально компактное пространство со счетной базой (в частности,  $X = \{x: -\infty < x < \infty\}$  — действительная прямая),  $\sigma$ -алгебра  $B_0$  порождается классом элементарных (т.е. конечномерных; см. [2]) открытых множеств. Известный метод Колмогорова задания случайного процесса его конечномерными распределениями (см. [3]) приводит к заданию именно бэровской меры. В пространстве  $U$  со счетной базой  $B_0$  совпадает с  $B$ , в пространстве с несчетной базой  $B \supset B_0$ . Если множество  $T$  индексов  $t$  несчетно, то вес пространства  $X^T$  будет несчетным и  $B \supset B_0$ . При этом многие

<sup>\*)</sup> Труды V Всесоюзного совещания по теории вероятностей и математической статистике (Ереван, 1958). — Ереван: Изд-во АН АрмССР, 1960. — С. 83–87.

интересные множества оказываются  $B$ -множествами. Поэтому приходится рассматривать те или иные продолжения бэровской меры. Вообще говоря, существует много таких продолжений. Однако если пространство  $U$  вполне регулярно, то может существовать главное продолжение бэровской меры до борелевской, которое определяется следующим условием “несчетной аддитивности”.

Для любой системы  $\{G_\alpha\}$  открытых множеств  $G_\alpha$

$$\mu\left(\bigcup_{\alpha} G_{\alpha}\right) = \sup_{k, (\alpha_1, \dots, \alpha_k)} \mu\left(\bigcup_{j=1}^k G_{\alpha_j}\right). \quad (2)$$

Можно записать условие (2) в двойственной форме (3).

Для любой системы  $\{F_\alpha\}$  замкнутых множеств  $F_\alpha$

$$\mu\left(\bigcap_{\alpha} F_{\alpha}\right) = \inf_{k, (\alpha_1, \dots, \alpha_k)} \mu\left(\bigcap_{j=1}^k F_{\alpha_j}\right). \quad (3)$$

Легко видеть, что продолжение меры по формуле (2) единственно, если только оно существует, и продолженная мера автоматически является топологической. Отметим, что вероятностная мера сепарабельного по Дубу (см. [4]) процесса является продолжением указанного типа. Именно действительная прямая пополняется точками  $-\infty$  и  $+\infty$  и рассматривается минимальная  $\sigma$ -алгебра  $D$ -множеств, содержащая все бэровские множества и все множества вида  $[a'_m, a''_m]^{(r'_m, r''_m)}$ , где  $[a'_m, a''_m]$  — отрезок  $a'_m \leq x \leq a''_m$  с рациональными концами и  $(r'_m, r''_m)$  — интервал  $r'_m < t < r''_m$  с рациональными концами. По определению полагаем, что

$$\mu\left\{[a'_m, a''_m]^{(r'_m, r''_m)}\right\} = \inf_{k; t_1, \dots, t_k \in (r'_m, r''_m)} \mu\{x(t): x_{t_j} \in [a'_m, a''_m], j = 1, \dots, k\}. \quad (4)$$

Можно доказать, что на  $\sigma$ -алгебре  $D$ ,  $B_0 \subset D \subset B$ , всегда существует единственная мера, совпадающая на  $B_0$  с продолжаемой бэровской и удовлетворяющая условию (4).

Всякая регулярная по Халмошу (см. [1]) мера на произвольной  $\sigma$ -алгебре  $S$  автоматически обладает свойством (1). Если пространство  $U$  — компакт (бикомпакт, по терминологии [2]), то верно и обратное. Регулярная мера обладает также свойством (2), если только  $G_\alpha \in S$  при всех  $\alpha$  и  $\bigcup_{\alpha} G_\alpha \in S$ . В [1] доказано, что всякая бэровская мера в компактном пространстве  $U$  может быть продолжена до регулярной борелевской меры, т.е. до борелевской меры, задаваемой условиями (2) и (3). Для некомпактного  $U$  легко построить пример, чтобы не существовало продолжения со свойством (2).

**ТЕОРЕМА 1.** *Если каждая конечная конечно-аддитивная топологическая мера  $\mu(\cdot)$ , заданная на алгебре  $S$ , содержащей открытую базу пространства  $U$ , может быть продолжена до несчетно-аддитивной борелевской меры, то пространство  $U$  компактно, и наоборот.*

Поэтому в случае некомпактного  $U$  можно применить следующую конструкцию. Пусть  $\beta U$  — чеховское расширение пространства  $U$ . Тогда мера  $\mu$  индуцирует в  $\beta U$  функцию множества  $\tilde{\mu}$  по правилу: для любого замкнутого  $F \in S$   $\tilde{\mu}\{[F]\} = \mu(F)$ , где  $[F]$  — замыкание  $F$  в  $\beta U$ .

ТЕОРЕМА 2. Каждая конечная конечно аддитивная топологическая мера  $\mu(\cdot)$ , заданная на алгебре  $S$ , содержащей открытую базу  $T_2$ -пространства  $U$ , индуцирует по правилу  $\tilde{\mu}\{[F]\} = \mu(F)$  несчетно-аддитивную топологическую меру на всех борелевских множествах пространства  $\beta U$ .

Если  $U$  является только  $T_1$ -пространством, то чеховское расширение  $\beta U$  в формулировке теоремы надо заменить уолменовским  $\omega U$ . Теорема остается справедливой, когда вместо  $\beta$  взято любое компактное расширение в  $U$  со счетной базой (если такое существует). В частности, распределение вероятностей случайного процесса, заданного по Колмогорову, всегда можно продолжить по формуле (2), если только добавить к фазовому пространству несобственные точки  $-\infty$  и  $+\infty$ .

При рассмотрении однородного по времени марковского процесса, задаваемого переходными вероятностями  $p(x; t, H)$  при некомпактном множестве состояний  $\Phi$ , приходится вводить несобственные мгновенные состояния. Возможен следующий метод пополнения  $\Phi$ . Будем смотреть на число  $p(x; t, H)$  как на координату точки  $x$  по оси с индексом  $(t, H)$ . Тогда каждая точка  $x \in \Phi$  отображается в некоторое подмножество  $V$  гиперкуба  $I^{(T, S)}$ , где  $I$  — единичный отрезок  $0 \leq p \leq 1$ ,  $T$  — множество моментов времени,  $S$  — класс всех измеримых множеств. Две точки  $x'$  и  $x''$  при этом склеиваются, если  $p(x'; t, H) \equiv p(x''; t, H)$  при всех  $t \in T$  и  $H \in S$ . Рассмотрим замыкание  $[V]$  в тихоновской топологии гиперкуба. Значения функций  $p(x; t, H)$  в добавленных точках определяются конструкцией. Очевидно также, что по аргументу  $H$  эти функции будут конечно-аддитивными топологическими мерами. Возьмем, далее, те  $H$ , которые в топологии гиперкуба являются замкнутыми в  $V$ , и положим  $\pi(x; t, [H]) = p(x; t, H)$ . По теореме 2 переходные вероятности  $\pi$  задают марковский процесс. К сожалению, при таком подходе вместе с действительно необходимыми мгновенными состояниями добавляются лишние; ср. [5].

Укажем еще одно применение теоремы 2. Пусть  $\nu(\cdot)$  — некоторая счетно-аддитивная мера, заданная на  $\sigma$ -алгебре  $S$ ,  $\zeta_\nu(\cdot)$  — случайная мера, заданная условием: каков бы ни был набор  $\{A_i\}_{i=1}^k$ ,  $k = 1, 2, \dots$ , случайные величины  $\zeta_\nu(A_i)$  имеют совместное нормальное распределение  $M\zeta_\nu(A_i) = 0$ ;  $M\zeta_\nu(A_i)\zeta_\nu(A_j) = \nu(A_i \cap A_j)$ . Это условие задает по Колмогорову некоторое распределение вероятностей в пространстве  $R^S$ , где  $R = (-\infty, +\infty)$ . Легко видеть, что  $P\{\zeta_\nu(A) = 0\} = 1$ , если  $\nu(A) = 0$ . Пополнив прямую  $R$  точками  $-\infty$  и  $+\infty$ , мы можем утверждать, что для регулярного продолжения

$$P\left\{\bigcap_{A \in S, \nu(A)=0} [\zeta_\nu(A) = 0]\right\} = 1.$$

Неправильно, однако, считать, что формула (2) дает возможность определить меры всех эффективно конструируемых множеств.

ОПРЕДЕЛЕНИЕ 2. Будем называть множество  $Q$   $V^*$ -измеримым, если  $Q$  принадлежит области определения каждой пополненной регулярной борелевской меры. В противоположном случае будем называть множество  $Q$   $V^*$ -неизмеримым.

$\sigma$ -алгебра  $V^*$ -измеримых множеств не исчерпывается одними  $V$ -множествами. К ней принадлежат, например,  $A$ -множества.

ТЕОРЕМА 3. В пространстве  $R^T$ , где  $R = [-\infty, +\infty]$ ,  $T = [0, 1]$ , являются  $B^*$ -неизмеримыми: а) все множества вида  $(a, b)^T$ , где  $(a, b)$  — интервал  $a < x < b$ ; б) множество  $N$  всех функций, непрерывных справа.

Отметим, что множество  $[a, b]^T$ , где  $[a, b]$  — отрезок, является замкнутым, и, тем самым,  $B$ -множеством. Для доказательства теоремы 3 конструируются две совпадающие на  $B$ -множествах меры,  $\mu_1(\cdot)$  и  $\mu_2(\cdot)$ , которые на исследуемом множестве  $Q$  принимают разные значения:  $\mu_1(Q) \neq \mu_2(Q)$ . Из теоремы 3 вытекает, что для продолжения меры могут понадобиться дополнительные соглашения помимо формулы (2).

### СПИСОК ЛИТЕРАТУРЫ

1. *Halmos P.R.* Measure theory. — N.Y., 1950.
2. *Александров П.С.* Введение в общую теорию множеств и функций. — М.-Л.: Гостехиздат, 1948.
3. *Колмогоров А.Н.* Основные понятия теории вероятностей. — М.-Л.: ОНТИ, 1936.
4. *Doob J.L.* Stochastic processes. — N.Y.-L., 1953.
5. *Chung K.L.* Proc. 3-rd Berkeley Symp. Math. Stat. and Prob. — Berkeley, 1956.

## ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ ДЛЯ НЕКОТОРЫХ КЛАССОВ СЛУЧАЙНЫХ ФУНКЦИЙ \*)

Пусть имеется последовательность случайных функций  $\xi_n(t)$  аргумента  $t \in T$  со значениями в  $\sigma$ -компактном топологическом пространстве  $X$ . Будем считать, что распределение вероятностей в пространстве  $X^T$ , соответствующее каждой случайной функции, определяется набором согласованных конечномерных распределений. Примем по определению  $\xi_n(\cdot) \rightrightarrows \xi_0(\cdot)$ , когда слабо сходятся соответствующие конечномерные распределения. Если  $\Phi$  — функционал, непрерывный в тихоновской топологии пространства  $X^T$ , то из  $\xi_n(\cdot) \rightrightarrows \xi_0(\cdot)$  следует, что случайная величина  $\Phi[\xi_n(\cdot)] \rightrightarrows \Phi[\xi_0(\cdot)]$ .

Пусть на множестве  $D$ , лежащем в пространстве  $X^T$ , введена некоторая своя топология  $\Theta$  и пусть  $\Psi$  — функционал, определенный на  $D$  и непрерывный в топологии  $\Theta$ . Простые примеры показывают, что условием  $P\{\xi_n(\cdot) \in D\} = 1$  при всех  $n \geq 0$  и  $\xi_n(\cdot) \rightrightarrows \xi_0(\cdot)$  может оказаться недостаточно для того, чтобы  $\Psi[\xi_n(\cdot)] \rightrightarrows \Psi[\xi_0(\cdot)]$ . Как показал Ю.В. Прохоров [1], в полном метрическом пространстве  $D$  для законности предельного перехода под знаком непрерывного функционала достаточны следующие условия:

- а)  $\xi_n(\cdot) \rightrightarrows \xi_0(\cdot)$ ;
- б) для любого  $\eta > 0$  существует такой  $\Theta$ -компакт  $D_\eta$ , что  $P\{\xi_n(\cdot) \in D_\eta\} > 1 - \eta$  при всех  $n \geq 0$ .

\*) Труды V Всесоюзного совещания по теории вероятностей и математической статистике (Ереван, 1958). — Ереван: Изд-во АН АрмССР, 1960. — С. 280–285.

Для пространства непрерывных функций с топологией равномерной сходимости условия компактности даются теоремой Арцела. Компактное множество непрерывных функций должно быть равномерно ограничено и равностепенно непрерывно. Поскольку непосредственно убедиться в компактности множества случайных траекторий бывает часто довольно затруднительно, приходится использовать достаточные критерии.

Пусть при каждом  $n \geq 0$  для случайного процесса  $\xi_n(\cdot)$  выполнен критерий непрерывности Колмогорова (см. [2]):

при всех  $t_1 < t_2$

$$M \|\xi_n(t_2) - \xi_n(t_1)\|^p < C |t_2 - t_1|^{1+r}, \quad (1)$$

где  $p > 0$ ,  $r > 0$ ,  $C > 0$  — константы, не зависящие от  $n$ .

Тогда можно оценить независимо от  $n$  распределение модуля непрерывности случайной траектории и показать, что со сколь угодно большой вероятностью имеет место равностепенная непрерывность. Равномерная ограниченность  $\xi_n(\cdot)$  следует из ограниченности по вероятности  $\xi_0(t)$  при фиксированном  $t$ , из условия  $\xi_n(\cdot) \Rightarrow \xi_0(\cdot)$  и из равностепенной непрерывности.

Критерий (1) может быть перенесен на случайные функции нескольких переменных, как это было сделано автором в [3]. При этом обычная разность  $\xi(t_2) - \xi(t_1)$  под знаком математического ожидания заменяется смешанной по всем переменным, а длина отрезка  $|t_2 - t_1|$  заменяется объемом параллелепипеда, по значениям случайных функций в вершинах которого строится смешанная разность. Для случая двух переменных этот критерий может быть сформулирован следующим образом.

а) При всех  $t_1 < t_2$  и  $s_1 < s_2$

$$M \|\xi_n(t_2, s_2) - \xi_n(t_1, s_2) - \xi_n(t_2, s_1) + \xi_n(t_1, s_1)\|^p < C |(t_2 - t_1)(s_2 - s_1)|^{1+r}, \quad (2)$$

где  $p > 0$ ,  $r > 0$ ,  $C > 0$  — константы, не зависящие от  $n$ .

б) Найдутся такие  $t_0$  и  $s_0$ , что при всех  $t_1 < t_2$  и  $s_1 < s_2$

$$\begin{aligned} M \|\xi_n(t_2, s_0) - \xi_n(t_1, s_0)\|^p &< C |t_2 - t_1|^{1+r}, \\ M \|\xi_n(t_0, s_2) - \xi_n(t_0, s_1)\|^p &< C |s_2 - s_1|^{1+r}. \end{aligned}$$

Условие (2) также позволяет оценить независимо от  $n$  распределение модуля непрерывности случайной функции; см. [3]. Таким образом, если  $\xi_n(\cdot, \cdot) \Rightarrow \xi_0(\cdot, \cdot)$  и выполнено (2) для любого функционала  $\Psi$ , непрерывного в равномерной топологии,  $\Psi[\xi_n(\cdot)] \Rightarrow \Psi[\xi_0(\cdot)]$ .

Рассмотрим теперь пространство функций без разрывов второго рода с топологией  $S$ , введенной А.В. Скороходом; см. [1, 4–6]. Условия компактности в этом пространстве получены А.Н. Колмогоровым.

Пусть функция  $f(t)$  принимает значения в каком-либо нормированном пространстве. Назовем *квазиколлебаием*  $f(t)$  на отрезке  $\Delta = [t_1, t_2]$  величину

$$\tilde{\omega}_f(\Delta) = \inf_{\substack{\theta=t^\pm \\ t_1 \leq t \leq t_2}} \max_{i=1,2} \omega_f(A_\theta^i),$$

где  $\omega_f(A_\theta^i)$  — колебание функции  $f(t)$  на множестве  $A_\theta^i$ :

$$\begin{aligned} A_\theta^1 &= [t_1, t], & A_\theta^2 &= (t, t_2] & \text{при } \theta = t^+; \\ A_\theta^1 &= [t_1, t), & A_\theta^2 &= [t, t_2] & \text{при } \theta = t^-. \end{aligned}$$

Соответственно назовем функцию  $\tilde{\omega}_f(\delta) = \sup_{\Delta; |\Delta| \leq \delta} \tilde{\omega}_f(\Delta)$  *модулем квазинепрерывности функции  $f$* . В этих терминах условие  $S$ -компактности напоминает условия Арцела: функции должны быть равномерно ограниченными и равномерно квазинепрерывными. В 1954 г., еще до создания общей теории, автор нашел, что условие:

при всех  $t_1 < t_2 < t_3$

$$M \|\xi_n(t_2) - \xi_n(t_1)\|^p \|\xi_n(t_3) - \xi_n(t_2)\|^q < C |t_3 - t_1|^{1+r}, \quad (3)$$

где  $p \geq 0$ ,  $q \geq 0$  и  $p + q > 0$ ,  $r > 0$ ,  $C > 0$  — константы, не зависящие от  $n$  делает законным предельный переход под знаком некоторых функционалов, если только траектории процессов  $\xi_n(t)$  не имеют разрывов второго рода. А.Н. Колмогоров предложил автору доказать, что условие (3) само обеспечивает с вероятностью единица отсутствие разрывов второго рода у траекторий  $\xi_n(t)$ . Соответствующее доказательство было опубликовано автором в [7]; и из проведенных в [7] выкладок нетрудно получить нужную оценку распределения модуля квазинепрерывности случайных траекторий. Таким образом, если  $\xi_n(\cdot) \Rightarrow \xi_0(\cdot)$  и выполнено (3), то для любого функционала  $\Psi$ , непрерывного в топологии  $S$ ,  $\Psi[\xi_n(\cdot)] \Rightarrow \Psi[\xi_0(\cdot)]$ . Этот результат также можно перенести на случайные функции нескольких переменных. Пусть функция  $f(t, s)$  задана на прямоугольнике  $T \times S$ . Будем рассматривать  $f(t, s)$  как функцию главного аргумента  $t$ , ставящую в соответствие каждому  $t = t_0$  элемент  $f(t_0, s)$  нормированного функционального пространства функций от  $s$  с нормой  $\|g(\cdot)\| = \sup_s \|g(s)\|$ . Тогда  $t$ -колебание функции  $f(t, s)$ , модуль  $t$ -квазинепрерывности, отсутствие  $t$ -разрывов второго рода вводятся согласно общим определениям. Рассмотрим пространство функций без разрыва второго рода по каждому переменному с обобщенной скороходовской топологией. Множество функций будет компактным в этой топологии тогда и только тогда, когда все функции равномерно ограничены и равномерно квазинепрерывны по каждому переменному. Соответствующее достаточное условие имеет вид:

при всех  $t_1 < t_2 < t_3$  и  $s' < s''$

$$\begin{aligned} M \left\| [\xi_n(t_2, s'') - \xi_n(t_2, s')] - [\xi_n(t_1, s'') - \xi_n(t_1, s')] \right\|^p \times \\ \times \left\| [\xi_n(t_3, s'') - \xi_n(t_3, s')] - [\xi_n(t_2, s'') - \xi_n(t_2, s')] \right\|^q < \\ < C |(t_3 - t_1)(s'' - s')|^{1+r}, \quad (4) \end{aligned}$$

при всех  $s_1 < s_2 < s_3$  и  $t' < t''$

$$\begin{aligned} M \left\| [\xi_n(t'', s_2) - \xi_n(t', s_2)] - [\xi_n(t'', s_1) - \xi_n(t', s_1)] \right\|^p \times \\ \times \left\| [\xi_n(t'', s_3) - \xi_n(t', s_3)] - [\xi_n(t'', s_2) - \xi_n(t', s_2)] \right\|^q < \\ < C |(t'' - t')(s_3 - s_1)|^{1+r}, \end{aligned}$$

где  $\xi_n(0, s) \equiv \xi_n(t, 0) \equiv \xi_n(0, 0)$ ,  $p \geq 0$ ,  $q \geq 0$  и  $p+q > 0$ ,  $r > 0$ ,  $C > 0$  — константы, не зависящие от  $n$ .

Можно рассматривать условия, промежуточные между (1) и (3). Пусть случайный процесс  $\xi(t)$  обладает свойством (3) при некоторых  $p > 0$  и  $q > 0$  и обладает в некоторой окрестности точки  $t_0$  свойством:

при всех  $t$  в окрестности  $t_0$

$$M \|\xi(t) - \xi(t_0)\|^u < Q |t - t_0|^v, \quad (5)$$

где  $u > 0$ ,  $v > 0$ ,  $Q > 0$  — константы.

Тогда все траектории процесса с вероятностью единица непрерывны в точке  $t_0$  и при любом

$$\gamma < \min \left( \frac{v}{u}, \frac{1+r}{p+q} \right)$$

и  $K > 0$  функция  $K |t - t_0|^\gamma$  является верхней функцией. Выполнение условия (5) менее ограничительно, чем выполнение условия (1) в точке  $t_0$ . Используя указанный вид верхней функции, можно доказать законность предельного перехода под знаком сингулярных функционалов от  $\xi_n(t)$ , обладающих свойствами (3) и (5) равномерно по  $n$ , типа

$$\int g \left( \frac{\xi_n(t) - \xi_n(t_0)}{|t - t_0|^\gamma}, t \right) dV(t),$$

где  $\int dV(t) < \infty$  и т.п.

Наконец, иногда полезно совсем слабое условие:

при всех  $t_1 < t_2 < t_3$

$$M \|\xi(t_2) - \xi(t_1)\|^p \|\xi(t_3) - \xi(t_2)\|^q < C |t_3 - t_1|^v, \quad (6)$$

где  $p \geq 0$ ,  $q \geq 0$  и  $p+q > 0$ ,  $v > 0$ ,  $C > 0$  — константы.

Если это условие выполнено, то каждое счетное всюду плотное множество можно принять за множество сепарабельности процесса.

Условия (1)–(6) удобно проверять для случайных функций, являющихся функциями распределения случайных мер. Тогда все разности в (1), (3), (5) и (6) превращаются в случайные меры отрезков, а смешанные разности в (2) и (4) — в случайные меры прямоугольников. В частности, пусть  $\eta_n(t) = \sqrt{n} [F_n^*(t) - t]$ ,  $F_n^*(x)$  — эмпирическая функция распределения по  $n$  наблюдениям,  $F(x)$  — непрерывная теоретическая функция распределения,  $t = F(x)$ ,  $0 \leq t \leq 1$ . Функция  $\eta_n(t)$  обладает свойством (3) с  $p = 2$ ,  $q = 2$ ,  $r = 1$ ,  $C = 1$  и свойством (5) с  $u = 2$ ,  $v = 1$ ,  $Q = 1$  в любой точке  $t_0$ . Как известно,  $\eta_n(t) \Rightarrow \zeta^0(t)$ , где  $\zeta^0(t)$  — условная винеровская траектория,  $\zeta^0(0) = \zeta^0(1) = 0$ . Поэтому условие (3) сводит вычисление асимптотического распределения непараметрического критерия согласия типа Колмогорова–Смирнова или Мизеса–Смирнова к вычислению распределения винеровского функционала. Условия (3) и (5) обосновывают подобное сведение для критериев типа критерия Андерсона и Дарлинга

$$n \int_{-\infty}^{+\infty} \frac{|F_n^*(x) - F(x)|^2}{F(x) [1 - F(x)]} dF(x)$$

(см. [8]).

В двумерном случае величина  $\eta_n(t, x) = \sqrt{n}[F_n^*(t, s) - ts]$ , где  $F_n^*(x, y)$  — эмпирическая функция распределения,  $F(x, y) = F(x, \infty)F(\infty, y)$  — непрерывная теоретическая функция распределения,  $t = F(x, \infty)$ ,  $s = F(\infty, y)$  слабо сходится к условному винеровскому полю  $\zeta^0(t, s)$  (см. [3]) с условием  $\zeta^0(1, 1) = 0$ . Поэтому, используя условие (4), можно легко вычислить асимптотическое распределение критерия независимости Хефдинга, указанное в [9] без доказательства.

#### СПИСОК ЛИТЕРАТУРЫ

1. Прохоров Ю.В. // Теория вероятн. и ее примен. 1956. Т. 1, вып. 2. С. 178–238.
2. Слуцкий Е.Е. // Giorn. Ist. Att. 1937. V. 8. P. 183–199.
3. Ченцов Н.Н. // Докл. АН СССР. 1956. Т. 106, № 4. С. 607–609.
4. Скороход А.В. // Докл. АН СССР. 1955. Т. 104, № 3. С. 364–367.
5. Колмогоров А.Н. // Теория вероятн. и ее примен. 1956. Т. 1, вып. 2. С. 239–247.
6. Скороход А.В. // Теория вероятн. и ее примен. 1956. Т. 1, вып. 3. С. 289–319.
7. Ченцов Н.Н. // Теория вероятн. и ее примен. 1956. Т. 1, вып. 1. С. 155–161.
8. Anderson T.W., Darling D.A. // Ann. Math. Stat. 1952. V. 23, № 2. P. 193–212.
9. Hoeffding W. // Ann. Math. Stat. 1948. V. 19. P. 546–557.

## ДУБОВСКИЕ МНОЖЕСТВА И ДУБОВСКИЕ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ<sup>\*)</sup>

В теории вероятностей принято задавать случайный процесс  $\xi(t)$ ,  $t \in T$ , его конечномерными борелевскими распределениями. Как доказал А.Н. Колмогоров [1], согласованные конечномерные распределения задают распределение вероятностей на некоторой  $\sigma$ -алгебре множеств. При этом многие интересные события, задаваемые ограничениями на поведение случайной траектории при несчетном числе моментов времени, оказываются неизмеримыми. Если предположить фазовое пространство  $R$  топологическим, то среди разных продолжений меры можно выделить главное продолжение, которое можно было бы назвать топологическим, как в [2], или, следуя Халмошу [3], называть его регулярным. Оно обладает следующим свойством “несчетной аддитивности”.

Для любой системы  $\{G_\alpha\}$  открытых в тихоновской топологии пространства  $R^T$  множеств  $G_\alpha$

$$\mu \left\{ \bigcup_{\alpha} G_{\alpha} \right\} = \sup_{k: (\alpha_1, \dots, \alpha_k)} \mu \left\{ \bigcup_{j=1}^k G_{\alpha_j} \right\}.$$

Краткое изложение этой конструкции было сделано автором в [2].

<sup>\*)</sup> Труды VI Всесоюзного совещания по теории вероятностей и математической статистике (Вильнюс, 1960). — Вильнюс: ГИПНЛ, 1962. — С. 483–492.

Более популярным является способ, предложенный в 30-х годах Дубом; см., например, [4]. Дуб вводит в двух вариантах понятие сепарабельного процесса (точнее, процесса с сепарабельными траекториями) и показывает, что при компактном фазовом пространстве каждый из вариантов гипотезы сепарабельности согласуется с любой системой согласованных конечномерных распределений. Таким образом, условие сепарабельности продолжает распределение вероятностей на некоторую  $\sigma$ -алгебру множеств, которые естественно называть дубовскими множествами.

В настоящей работе дается обобщение конструкции Дуба, позволяющее отказаться почти полностью от топологических ограничений на фазовое пространство.

Введем следующие обозначения. Пусть  $R$  — фазовое пространство случайного процесса,  $T$  — множество аргументов. Реализациями случайного процесса будут всевозможные функции  $x(t)$  аргумента  $t \in T$  со значениями в  $R$ . Каждую такую функцию можно рассматривать как точку пространства  $R^T$ , т.е. как набор координат  $\{x_t\}$ ,  $x_t = x(t)$ , где номера координат занумерованы всеми индексами  $t \in T$ .

Пусть  $Q$  — некоторая  $\sigma$ -алгебра множеств фазового пространства  $R$  и  $K(Q)$  — алгебра подмножеств пространства  $R^T$ , порожденная одномерными цилиндрическими множествами, т.е. минимальная  $\sigma$ -алгебра, содержащая все цилиндрические множества вида  $a_t \times R^{T-t}$ , где  $a_t \in Q$ .

Предположим, что на  $K(Q)$  задана вероятностная мера; попытаемся ее продолжить на некоторые специальные подмножества  $R^T$ .

**ОПРЕДЕЛЕНИЕ 1.** Будем называть параллелепипед  $A$  пространства  $R^T$ ,  $A = \times_{t \in T} a_t$ , *коридором*, если среди множеств  $\pi_t A = a_t$ ,  $a_t \subset R$ , проекций параллелепипеда  $A$  на координатные оси  $R_t$  имеется лишь конечное число отличных друг от друга.

**ОПРЕДЕЛЕНИЕ 2.** Непересекающиеся подмножества  $T_1, \dots, T_k$  множества индексов  $T$  мы будем называть *множествами постоянства* для коридора  $A$ , если для любого  $T_i$  из того, что  $t' \in T_i$  и  $t'' \in T$ , следует, что  $a_{t'} = a_{t''}$ .

Заметим, что в определении 2 не требуется, чтобы система  $T_1, \dots, T_k$  была минимальной, так что любое конечное разбиение множеств постоянства снова образует систему множеств постоянства.

**СЛЕДСТВИЕ.** *Пересечение двух или конечного числа коридоров является снова коридором.*

**ОПРЕДЕЛЕНИЕ 3.** Пусть  $\Gamma$  — конечный или счетный набор индексов  $t$ ,  $t \in T$ . Назовем параллелепипед  $A_\Gamma$   $\Gamma$ -аналогом параллелепипеда, в частности, коридора,  $A = \times_{t \in T} a_t$ , если  $A_\Gamma = (\times_{t \in \Gamma} a_t) \times R^{T-\Gamma}$ .

**ЗАМЕЧАНИЕ 1.** Для любого  $\Gamma$   $\Gamma$ -аналог коридора является коридором.

**ОПРЕДЕЛЕНИЕ 4.** Будем говорить, что некоторый класс  $F$  подмножеств пространства  $R$ , замкнутый относительно пересечения, *образует в  $R$  компактную систему*, если всякая монотонно убывающая счетная последовательность непустых множеств из  $F$  имеет непустое пересечение.

**ЗАМЕЧАНИЕ 2.** Структура, порожденная компактной системой, в силу дистрибутивности операций объединения и пересечения сама является компактной системой.

Примерами компактных систем могут служить совокупность всех отрезков на прямой или класс всех замкнутых множеств на прямой, пополненной точками  $+\infty$  и  $-\infty$ .

**ТЕОРЕМА 1.** Пусть  $K$  — произвольный счетный подкласс класса  $K(Q)$  и пусть  $H$  — произвольный счетный класс коридоров с проекциями из произвольной компактной системы  $F$ ,  $F \subset Q$ .

Тогда существует счетное множество  $\Gamma_0$  индексов  $t \in T$  такое, что при любом  $\Gamma \supseteq \Gamma_0$  существует гомоморфизм  $\sigma$ -алгебры  $S$ , порожденной классом  $K \cup H$ , на  $\sigma$ -алгебру  $S_\Gamma$ , порожденную классом  $K \cup H_\Gamma$ , состоящим из  $\Gamma$ -аналогов коридоров класса  $H$  и множеств класса  $K$ .

**Доказательство.** Известно (см. [3]) следующее предложение:  $\sigma$ -алгебра множеств  $S(L)$ , порожденная классом  $L$ , совпадает с объединением  $\sigma$ -алгебр  $S(L_\beta)$ , порожденных всеми счетными подклассами  $L_\beta$  класса  $L$ .

Следовательно, каждое множество из  $K(Q)$  порождено счетным числом одномерных цилиндрических множеств вида  $a_t \times R^{T-t}$ , и мы, не ограничивая общности, можем считать, что подкласс  $K$  состоит только из множеств такого вида. Заметим, что такие множества тоже являются коридорами, только  $a_t$  не обязательно входит в систему  $F$ . Занумеруем все коридоры обоих классов общей нумерацией  $C_1, C_2, C_3, \dots$ . Пусть  $T_1^{(1)}, \dots, T_{k(1)}^{(1)}$  — разбиение множества индексов  $T$  на максимальные множества постоянства для  $C_1$ . Пусть  $T_1^{(2)}, \dots, T_{k(2)}^{(2)}$  — разбиение  $T$  на максимальные общие множества постоянства для  $C_1$  и  $C_2$  и, далее,  $T_1^{(m)}, \dots, T_{k(m)}^{(m)}$  — разбиение  $T$  на максимальные общие множества постоянства для  $C_1, C_2, \dots, C_m$ . Очевидно, каждое разбиение получается подразделением множеств предыдущего разбиения. Построим набор  $\Gamma_0$  следующим образом. Выберем в каждом из множеств  $T_i^{(1)}$  по точке и занумеруем их  $t_1, \dots, t_{k(1)}$ . Эти точки попадают в некоторые из множеств  $T_i^{(2)}$ , причем в каждое множество не более одной точки. Выберем в тех из множеств  $T_i^{(2)}$ , куда точки  $t_1, \dots, t_{k(1)}$  не попали, по точке и занумеруем их  $t_{k(1)+1}, \dots, t_{k(2)}$ . Прделаем такую же операцию при всех  $m$ , выбирая каждый раз по точке из тех множеств разбиения, куда не попали точки, выбранные на предыдущих шагах. Полученную последовательность точек  $\{t_i\}$  обозначим через  $\Gamma_0$ .

Условимся записывать  $C^1 = C$  и  $C^{-1} = \overline{C}$ . Тогда имеет место следующая

**ЛЕММА.** Пусть  $\Gamma \supseteq \Gamma_0$ . Если для некоторой последовательности  $\{\alpha_i\}_{i=1}^\infty$ ,  $\alpha_i = \pm 1$ ,  $\bigcap_{i=1}^\infty C_i^{\alpha_i} = \emptyset$ , то и

$$\bigcap_{i=1}^\infty (C_i)_\Gamma^{\alpha_i} = \emptyset,$$

где  $(C_i)_\Gamma$  —  $\Gamma$ -аналоги коридоров  $C_i$ .

Доказательство леммы будет проведено отдельно; оно составляет самую тонкую часть наших рассуждений. А сейчас мы воспользуемся ею, чтобы построить гомоморфизм алгебры коридоров в алгебру их  $\Gamma$ -аналогов при условии, что  $\Gamma$  содержит построенное нами выше множество индексов  $\Gamma_0$ .

Рассмотрим всевозможные последовательности  $\{\alpha_i\}_{i=1}^\infty$ ,  $\alpha_i = \pm 1$ , и поставим в соответствие каждой последовательности  $\alpha$  множество  $D_\alpha = \bigcap_{i=1}^\infty C_i^{\alpha_i}$ . Тогда множества, представимые в виде  $D_A = \bigcup_{\alpha \in A} D_\alpha$ , образуют булеву алгебру  $B$ ,

содержащую  $\sigma$ -алгебру  $S$ , порожденную множествами  $C_i$ . В самом деле,  $\bigcup_i D_{A_i} = \bigcup_{\alpha \in A} D_\alpha$ ,  $A = \bigcup_i A_i$ . Далее,

$$D_\alpha \cap D_\beta = \begin{cases} \emptyset & \text{при } \alpha \neq \beta, \\ D_\alpha & \text{при } \alpha = \beta. \end{cases}$$

Поэтому  $\bigcap_i D_{A_i} = \bigcup_{\alpha \in A} D_\alpha = D_A$ ,  $A = \bigcap_i A_i$ . Наконец,  $D_{\bar{A}} = \overline{D_A}$ . Таким образом, множества  $D_A$  образуют  $\sigma$ -алгебру, которая во всяком случае не уже  $\sigma$ -алгебры  $S$  — наименьшей  $\sigma$ -алгебры, содержащей множества  $C_i$ . Алгебра  $I$  всех подмножеств множества всех последовательностей  $\alpha$  является свободной булевой алгеброй. Соответствие  $A \rightarrow D_A$  задает гомоморфизм алгебры  $I$  на алгебру  $B$ .

Так как среди множеств  $D_\alpha$  могут оказаться пустые (для этого достаточно, например, чтобы  $C_1 \cap C_2 \neq \emptyset$  или  $C_1 \cap \overline{C_2} = \emptyset$ ), то каждому множеству из  $B$  будет соответствовать, вообще говоря, несколько записей в виде  $D_A$ . В самом деле, если множество индексов  $\alpha$  пустых  $D_\alpha$  мы обозначим через  $\Lambda$ , то  $D_A = D_\Lambda \bigcup D_{A_1}$ , если  $A_1 \subseteq \Lambda$ . Множество  $\Lambda$  образует идеал алгебры  $I$ , и можно показать, что  $B \simeq I/\Lambda$ . Построим теперь точно такую же алгебру  $B_\Gamma$  по множествам  $(C_i)_\Gamma$ . Точно так же множество индексов  $\alpha$ , соответствующих пустым  $D_\alpha^\Gamma$ , будет образовывать идеал  $\Lambda_\Gamma$  алгебры  $I$  и  $B_\Gamma \simeq I/\Lambda_\Gamma$ . Так как  $\Gamma \supseteq \Gamma_0$ , то по лемме из  $D_\alpha = \emptyset$  следует  $D_\alpha^\Gamma = \emptyset$ , т.е.  $\Lambda \subseteq \Lambda_\Gamma$ . Нам остается лишь провести рассуждения, показывающие, что  $B_\Gamma \cong B / \frac{\Lambda}{\Lambda_\Gamma}$ . Поставим в соответствие каждо-

му множеству  $D_A$  его аналог — множество  $D_A^\Gamma$ . Покажем, что это соответствие порождает гомоморфизм  $S$  в  $S_\Gamma$ . В самом деле, каждой записи множества из  $S$  мы поставили множество из  $B_\Gamma$ . Пусть  $D_A \supseteq D_A$ . Тогда множество  $A$  можно разбить на  $A_2$  и  $A_3$  так, что  $A_3 \subseteq A_1$  (например,  $A_3 = A \cap A_1$ ),  $A_2 \subseteq \Lambda$ ,  $A_2 \cap A_3 = \emptyset$ , откуда  $D_A = D_{A_3} \cup D_{A_2}$ , где  $D_{A_2} = \emptyset$ . Но раз  $A_3 \subseteq A_1$ , то и  $D_{A_3}^\Gamma \subseteq D_{A_1}^\Gamma$ , а раз  $A_2 \subseteq \Lambda \subseteq \Lambda_\Gamma$ , то и  $D_{A_2}^\Gamma = \emptyset$ . Поэтому  $D_{A_2}^\Gamma \supseteq D_{A_1}^\Gamma$ . Таким образом, каждой записи множества из  $B$  соответствует одно и то же множество из  $B_\Gamma$ , так как из  $D_{A_1} = D_{A_2}$  следует  $D_{A_1}^\Gamma \supseteq D_{A_2}^\Gamma$  и  $D_{A_1}^\Gamma \subseteq D_{A_2}^\Gamma$ . Очевидно, объединению множеств соответствует объединение, пересечению — пересечение, дополнению — дополнение так, как это имеет место для записей. Далее, множеству  $C_j \bigcup_{\alpha: \alpha_j=1} D_\alpha$  соответствует множество  $(C_j)_\Gamma = \bigcup_{\alpha: \alpha_j=1} D_\alpha^\Gamma$ . Наконец, очевидно, что множества из  $B$ , отображаемые в множества из  $S_\Gamma$ , образуют  $\sigma$ -алгебру. Так как эта  $\sigma$ -алгебра содержит класс  $C_i$ , то она во всяком случае содержит и порожденную множествами  $C_i$   $\sigma$ -алгебру  $S$ . Этим теорема доказана.

**ЗАМЕЧАНИЕ 3.** Из доказательства теоремы следует, что мы можем называть множество  $V_\Gamma$   $\Gamma$ -аналогом множества  $V$ , если  $V_\Gamma$  получается из  $(C_i)_\Gamma$  теми же операциями, что и  $V$  из  $C_i$ . Точнее,  $V_\Gamma$  является  $\Gamma$ -аналогом множества  $V$ , если  $V_\Gamma = \bigcup_{\alpha \in A(V)} D_\alpha^\Gamma$  и  $V = \bigcup_{\alpha \in A(V)} D_\alpha$ , где  $\alpha = \{\alpha_i\}_{i=1}^\infty$ ,  $\alpha_i = \pm 1$ ;  $D_\alpha = \bigcap_{i=1}^\infty C_i^{\alpha_i}$ ;  $D_\alpha^\Gamma = \bigcap_{i=1}^\infty (C_i)_\Gamma^{\alpha_i}$ . Это определение не зависит от способа представления  $V$  в виде  $\bigcup_\alpha D_\alpha$ , во всяком случае, если  $\Gamma \supseteq \Gamma_0$ , где  $\Gamma_0$  — построенное при доказательстве специальное множество индексов  $t$ .

**Доказательство леммы.** Покажем, что неравенство  $\bigcap_{i=1}^\infty (C_i)_\Gamma^{\alpha_i} \neq \emptyset$  влечет за собой, что  $\bigcap_{i=1}^\infty C_i^{\alpha_i} \neq \emptyset$ , откуда утверждение леммы следует тривиаль-

ным образом.

Пусть существует точка  $\{x_t\} \in \bigcap_{i=1}^{\infty} (C_i)_{\Gamma}^{\alpha_i} \neq \emptyset$ . Тогда  $\{x_t\}$  можно подправить до  $\{\tilde{x}_t\}$ , так что  $x_t = \tilde{x}_t$  при  $t \in \Gamma$  и  $\{\tilde{x}_t\} \in \bigcap_{i=1}^{\infty} C_i^{\alpha_i}$ . В самом деле, положим  $\tilde{x}_t = x_t$ , при  $t \in \Gamma$ . Пусть  $t \notin \Gamma$ . Тогда существует бесконечная последовательность точек  $\{t_{\beta(i)}\}_{i=1}^{\infty}$ ,  $t_{\beta(i)} \in \Gamma_0$ , таких, что при каждом  $i$   $t$  и  $t_{\beta(i)}$  принадлежат одному и тому же множеству разбиения  $T_{\gamma(i)}^{(i)}$ . Рассмотрим проекции  $c_{\gamma(i)}$  коридоров  $C_i$ , соответствующих множествам разбиения  $T_{\gamma(i)}^{(i)}$ . Очевидно,  $c_{\gamma(i)}$  принадлежат к системе  $F$ , так как сечения, не входящие в  $F$ , могли быть только у коридоров класса  $K$ , но соответствующие множества постоянства в этом случае состояли ровно из одной точки, которая поэтому автоматически принадлежит  $\Gamma_0$ , а стало быть, и  $\Gamma$ . Обозначим  $d_m = \bigcap_{i: \alpha_i=1} c_{\gamma(i)}$ . Очевидно,  $d_m \in F$ . Так как  $\{x_t\} \in \bigcap_i (C_i)_{\Gamma}^{\alpha_i}$ , то  $\{x_t\} \in \bigcap_{i: \alpha_i=1} (C_i)_{\Gamma}$ , поэтому  $x(t_{\beta(i)}) \in d_m$ , т.е.  $d_m$  не пусто. Так как  $d_m$  принадлежат к системе  $F$ , то  $\bigcap_m d_m \neq \emptyset$ ; и мы можем положить  $\tilde{x}_t \in \bigcap_m d_m$ . Точно такое же построение осуществим при всех  $t \in T - \Gamma$ . Остается проверить, что  $\{\tilde{x}_t\} \in \bigcap_i C_i^{\alpha_i}$ , т.е.  $\{\tilde{x}_t\} \in C_i^{\alpha_i}$  при любом  $i$ . Пусть  $\alpha_i = -1$ . Тогда  $\{\tilde{x}_t\} \notin (C_i)_{\Gamma}$ , т.е. существует такое  $t_i \in \Gamma$ , что  $x_{t_i} \notin b_i$ , где  $b_i = \pi_{t_i} C_i$ . Так как  $\tilde{x}_{t_i} = x_{t_i}$ , то  $x_{t_i} \notin \pi_{t_i} C_i$ , и, следовательно,  $\{\tilde{x}\} \notin C_i$ , т.е.  $\{\tilde{x}\} \in C_i^{-1}$ . Пусть теперь  $\alpha_i = +1$ . Если  $t \in \Gamma$ , то  $\tilde{x}_t \in \pi_t C_i$ , так как  $\{x_t\} \in \bigcap_i (C_i)_{\Gamma}^{\alpha_i}$ . Если  $t \notin \Gamma$ , то по построению  $\tilde{x}_t \in d_i$  и, следовательно,  $\tilde{x}_t \in \pi_t C_i$ , где  $\pi_t C$  обозначает всюду проекцию коридора  $C$  на ось  $R_t$ . Таким образом,  $\{\tilde{x}_k\} \in C_{\Gamma}^{\alpha_i}$  и  $\{\tilde{x}\} \in \bigcap_i C_i^{\alpha_i}$ . Лемма доказана.

**ТЕОРЕМА 2.** Пусть на  $\sigma$ -алгебре  $K(Q)$  задана конечная счетно-аддитивная мера  $\mu(\cdot)$  и пусть  $H$  — класс коридоров с проекциями, входящими в некоторую компактную систему  $F$ ,  $F \subset Q$ .

Тогда мера  $\mu$  продолжается до меры  $\tilde{\mu}$  на  $\sigma$ -алгебре  $S(Q, H)$ , порожденной объединением классов  $K(Q)$  и  $H$ , причем мера коридоров будет даваться формулой

$$\tilde{\mu}(A) = \lim_{\Gamma} \mu(A_{\Gamma}),$$

где  $\lim_{\Gamma}$  понимается как предел по частично упорядоченному множеству  $\Gamma_1 > \Gamma_2$ , если  $\Gamma_1 \supset \Gamma_2$ .

Если воспользоваться данным в замечании 3 расширением определения  $\Gamma$ -аналога, то эта же формула будет задавать меру любого множества из  $S(Q, H)$ .

**Доказательство.** Для каждого коридора  $C$  из  $\Gamma_1 \supset \Gamma_2$  следует  $C_{\Gamma_1} \subset C_{\Gamma_2}$ . Поэтому для каждого  $C$   $\inf \mu(C_{\Gamma})$  существует и достигается на некотором счетном  $\Gamma = \Gamma(C)$ , а значит, и на всех  $\Gamma \supseteq \Gamma(C)$ . Таким образом, мы можем положить  $\tilde{\mu}(C) = \lim_{\Gamma} \mu(C_{\Gamma})$ . Остается показать, что так определенную функцию множества можно продолжить до счетно-аддитивной функции на  $S$ . Пусть  $V \in S$ . Существует счетный класс  $L$  коридоров  $C_i$  таких, что  $V \in S(L)$ , причем либо  $C_i \in H$ , либо  $C_i$  — одномерное цилиндрическое множество. Пусть  $\Gamma_0$  — множество индексов  $t$  для класса коридоров  $L$ , построенное по теореме 1. Выбираем множества  $\Gamma_i$  так, чтобы  $\mu(C_i)_{\Gamma_i} = \tilde{\mu}(C_i)$ , кладем  $\Gamma \supseteq \bigcup_{i=0}^{\infty} \Gamma_i$  и присваиваем множеству  $V$  меру его  $\Gamma$ -аналога (см. замечание 3).

Как было отмечено, определение  $\Gamma$ -аналога (при  $\Gamma \supseteq \Gamma_0$ ) не зависит от записи  $V$ . Поэтому мера  $\tilde{\mu}(V)$  не зависит от способа записи  $V$ , хотя, быть может,

зависит от  $\Gamma$  и  $L$ . Покажем, что при фиксированных  $\Gamma$  и  $L$  получившаяся функция множества счетно-аддитивна. Пусть  $V_i \in S(L_\Gamma)$  и  $V_i \cap V_j = \emptyset, i \neq j; i, j = 1, 2, \dots$ . Тогда  $(V_i)_\Gamma \in S(L_\Gamma)$  и по лемме  $(V_i)_\Gamma \cap (V_j)_\Gamma = \emptyset; i \neq j; i, j = 1, 2, \dots$ . Наконец,  $(\bigcup_i V_i)_\Gamma = \bigcup_i (V_i)_\Gamma$ . Поэтому  $\tilde{\mu}_{L,\Gamma}(\bigcup_i V_i) = \mu(\bigcup_i (V_i)_\Gamma) = \sum_i \mu(V_i)_\Gamma = \sum_i \tilde{\mu}_{L,\Gamma}(V_i)$ . Далее, при  $\Gamma \supseteq \Gamma_i$   $\mu_{L,\Gamma}(C_i) = \mu(C_i)_\Gamma = \tilde{\mu}(C_i)$ . Так как на  $\sigma$ -алгебре две меры совпадают, если они совпадают на порождающих множествах, то введенная нами мера  $\tilde{\mu}_{L,\Gamma}$  не зависит от  $\Gamma \supseteq \bigcup_{i=0}^\infty \Gamma_i$  и  $L$ . Наконец, мера  $\mu$  счетно-аддитивна, так как для любой счетной последовательности  $V_i \in S(L_i)$  имеем  $V_i \in S(L)$ , где  $L = \bigcup_{i=1}^\infty L_i$ , а для фиксированного  $L$  счетная аддитивность уже доказана.

**СЛЕДСТВИЕ.** Если множество  $V$  принадлежит к  $\sigma$ -структуре, порожденной коридорами класса  $H$ , то его  $\tilde{\mu}$ -мера определяется по формуле  $\tilde{\mu}(V) = \min_\Gamma \mu(V_\Gamma)$ .

Теорема 2 показывает, что, оставаясь в классе событий из  $S(Q, H)$ , мы можем считать все реализации случайного процесса сепарабельными и вычислять вероятности событий этого класса из условий сепарабельности. Что же касается общего для всех событий множества сепарабельности  $\Gamma_0$ , то оно существует только лишь, когда  $H$  порождено счетным классом коридоров (хотя само при этом не обязано быть счетным).

Выделим некоторый специальный класс коридоров.

**ОПРЕДЕЛЕНИЕ 5.** Назовем класс  $H$  коридоров пространства  $R^T$  классом  $Z^Y$ , если он состоит из всех возможных коридоров, множества постоянства которых принадлежат алгебре  $Y$  подмножеств множества индексов  $T$ , а сечения принадлежат компактной системе  $Z$  фазового пространства  $R$ . Через  $S(Z^Y)$  будем обозначать  $\sigma$ -алгебру, порожденную классом  $Z^Y$ .

**ЗАМЕЧАНИЕ 4.** Всякий класс  $Z^Y$  замкнут относительно операции пересечения и является компактной системой в пространстве  $R^T$ .

Применим теперь наши результаты к классическому случаю:  $R$  — пополненная числовая прямая,  $T$  — действительная ось или ее отрезок. Под  $\sigma$ -алгеброй  $Q$  мы будем понимать всю  $\sigma$ -алгебру борелевских множеств на прямой, а под  $K = K(Q)$  — алгебру порожденных ими колмогоровских множеств.

**ТЕОРЕМА 3.** Пусть  $Y$  — алгебра множеств, порожденная всеми интервалами оси  $T$ ,  $Z_1$  состоит из всех отрезков,  $Z_2$  состоит из всех конечных объединений отрезков,  $Z_3$  состоит из всех замкнутых множеств пополненной прямой  $R$ . Тогда:

1)  $K(Q) \subset S(Z_1^Y) \subset S(Z_2^Y) = S(Z_3^Y) \subset S(Z_1^{Y_1})$ , где  $Y_1$  — алгебра всех подмножеств  $T$ ;

2)  $\sigma$ -алгебры  $S(Z_1^Y)$  и  $S(Z_3^Y)$  могут быть получены из  $\sigma$ -алгебры  $K(Q)$  всех колмогоровских множеств добавлением счетного числа образующих.

Эти  $\sigma$ -алгебры  $D_1 = S(Z_1^Y)$  и  $D = S(Z_2^Y) = S(Z_3^Y)$  естественно называть алгебрами дубовских множеств (в широком и узком смысле слова), а продолженное на них по формуле теоремы 2 распределение вероятностей — дубовским распределением. Общее множество сепарабельности  $\Gamma_0$ , которое существует в силу утверждения 2), зависит от конечномерных распределений.

Теорема 3 может быть почти дословно перенесена на случай, когда  $R$  и  $T$  — евклидовы пространства произвольной размерности, и даже на случай, когда  $R$  — пополненное точкой  $+\infty$  метрическое пространство со счетной базой. Од-

нако во всех этих случаях все коридоры класса  $Z^{Y_1}$  являются замкнутыми в тихоновской топологии пространства  $R^T$ , так как системы  $Z$  содержат только замкнутые подмножества компактного пространства  $R$ . Поэтому классы  $S(Z^Y)$ , даже  $S(Z_3^{Y_1})$ , входят в  $\sigma$ -алгебру всех борелевских множеств компактного пространства  $R^T$ , и продолжение по Дубу является частным случаем продолжения бэрвской меры до борелевской (см. [2]).

**Доказательство.** Докажем сперва, что  $S(Z_2^Y) = S(Z_3^Y)$ . Так как во всяком случае  $S(Z_2^Y) \subseteq S(Z_3^Y)$ , то нам достаточно показать, что каждый коридор  $A$ ,  $A \in Z_3^Y$ , представим в виде  $A = \bigcap_{i=1}^{\infty} A_i$ , где  $A_i \in Z_2^Y$ . Пусть множества постоянства коридора  $A$  суть  $T_1, \dots, T_k$ , а соответствующие проекции —  $c_1, \dots, c_k$ . Всякое замкнутое множество  $c_j$  на пополненной прямой представимо в виде  $c_j = \bigcap_{i=1}^{\infty} f_{ij}$ , где  $f_{ij}$  — объединения конечного числа отрезков. Отсюда искомые  $A_i$  можно записать в виде  $A_i = \bigtimes_{j=1}^k (f_{ij})^{T_j}$ .

Докажем теперь утверждение 2). В качестве счетной системы коридоров выберем для  $D_1$  коридоры вида  $R^{T-(a_j, b_j)} \times [c_k, d_k]^{(a_j, b_j)}$ , где  $(a_j, b_j)$  — интервалы с рациональными концами,  $[c_k, d_k]$  — отрезки с рациональными концами и все рациональные интервалы и отрезки занумерованы. Покажем, что любой коридор из  $D_1$  является пересечением не более счетного числа коридоров указанного вида и еще, быть может, коридоров из алгебры  $K$ . Поскольку каждый коридор является пересечением конечного числа коридоров вида  $R^{T-T_0} \times F^{T_0}$ , где  $T_0$  — либо точка, либо отрезок, либо полуинтервал, либо интервал, а  $F$  — отрезок, то надо показать утверждение лишь для них. Это же почти очевидно.

В случае  $D$  выберем коридоры вида  $R^{T-(a_j, b_j)} \times F_k^{(a_j, b_j)}$ , где через  $F_k$  занумерованы все конечные объединения отрезков с рациональными концами.

Для доказательства утверждения, что данное множество  $A \in S_2$  не принадлежит  $S_1$ ,  $S_1 \subseteq S_2$ , и, следовательно,  $S_2 \supset S_1$ , мы будем строить на  $S_2$  две меры,  $\mu_1$  и  $\mu_2$ ;  $\mu_1(A) \neq \mu_2(A)$ , но совпадающие на  $S_1$  всюду.

Покажем, что  $K \neq D_1$ . Пусть  $\{\omega\} = \Omega = [0, 1]$  с лебеговой длиной в качестве меры. Предположим для простоты, что  $T$  также есть  $[0, 1]$ . Построим измеримые отображения  $\xi_+(t)$  и  $\xi_0(t)$  отрезка  $\Omega$  в пространстве  $R^T$ :

$$\begin{aligned} \xi_+(t; \omega) &= 0 \quad \text{при } \omega \neq t, \\ \xi_+(t; \omega) &= 1 \quad \text{при } \omega = t. \\ \xi_0(t; \omega) &\equiv 0. \end{aligned}$$

Очевидно, все коридоры вида  $R^{T-T_0} \times F^{T_0}$  (коридоры с одной ступенью) измеримы как при  $T_0 \in Y$ , так и когда  $T_0$  состоит из одной точки. Поэтому все коридоры измеримы и все  $D_1$ -множества также измеримы. Поскольку для фиксированного  $t$   $P\{\xi_0(t) = 0\} = 1 = P\{\xi_+(t) = 0\}$ , то  $\mu_1 = \mu_2$  на всех колмогоровских множествах, так как они совпадают на всех порождающих алгебру  $K$  одномерных цилиндрических множествах. А для коридора  $C = [-1/2, +1/2]^{[0, 1]}$  меры различаются:

$$\begin{aligned} \mu_1(C) &= P\left\{ -\frac{1}{2} \leq \xi_0(t) \leq \frac{1}{2} \quad \text{при всех } t \right\} = 1, \\ \mu_2(C) &= P\left\{ -\frac{1}{2} \leq \xi_+(t) \leq \frac{1}{2} \quad \text{при всех } t \right\} = 0, \end{aligned}$$

что и требовалось.

Совершенно аналогично с помощью двух случайных функций от того же аргумента

$$\begin{aligned} \xi_+(t, \omega) &= 0 \quad \text{при } \omega < t, & \xi_0(t, \omega) &= 0 \quad \text{при } \omega < t, \\ \xi_+(t, \omega) &= 1 \quad \text{при } \omega \geq t, & \xi_0(t, \omega) &= \frac{1}{2} \quad \text{при } \omega = t, \\ & & \xi_0(t, \omega) &= 1 \quad \text{при } \omega > t \end{aligned}$$

и коридора  $C$  с сечением  $[-1/3, +1/3] \cup [2/3, 4/3]$  доказывается, что  $D \supset D_1$ .

Наконец, подобным же, но более хитрым примером можно показать, что  $S(Z_3^{Y_1}) \supset D$ , и, тем самым,  $\sigma$ -алгебра  $D$  всех дубовских множеств уже  $\sigma$ -алгебры  $B$  всех борелевских подмножеств  $R^T$ . Проще, однако, сослаться на теорему 4, где утверждается, что каждое непустое  $D$ -множество содержит функцию второго класса по классификации Бэра. Так как компактное множество  $C$  класса  $S(Z_3^{Y_1})$ , состоящее из одной лишь функции строго третьего класса, таковой не содержит, то  $B \supseteq S(Z_3^{Y_1}) \supset D$ .

В работе [5] показано, что внешняя колмогоровская мера множества всех бэровских функций второго класса равна 1, и поэтому непротиворечиво считать все случайные траектории функциями второго класса. Докажем аналогичное утверждение для дубовской меры.

**ТЕОРЕМА 4.** *Всякое непустое  $D$ -множество содержит хотя бы одну бэровскую функцию второго класса.*

**СЛЕДСТВИЕ.** *Множество всех бэровских функций второго класса всегда имеет внешнюю  $D$ -меру, равную 1.*

**Доказательство.** Пусть  $H \in D$ . Тогда  $H$  принадлежит  $\sigma$ -алгебре, порожденной некоторым классом одноступенчатых коридоров, т.е. коридоров вида  $K_i = R^{T-t_i} \times E_i^{t_i}$  или  $C_j = R^{T-T_j} \times F^{T_j}$ , где  $F$  — объединение конечного числа отрезков с рациональными концами,  $E$  — произвольное борелевское подмножество  $R$ , а  $T_j$  — интервалы с рациональными концами (см. доказательство теоремы 3). При доказательстве теоремы 1 (см. замечание 3) мы доказали, что всякий элемент  $\sigma$ -алгебры может быть записан через образующие в виде

$$H = \bigcup_{(\alpha, \beta) \in N(H)} \left\{ \left[ \bigcap_{i=1}^{\infty} K_i^{\alpha_i} \right] \cap \left[ \bigcap_{j=1}^{\infty} C_j^{\beta_j} \right] \right\},$$

где  $\alpha = \{\alpha_i\}_{i=1}^{\infty}$ ,  $\alpha_i = -1$  или  $\alpha_i = +1$ ;  $\beta = \{\beta_j\}_{j=1}^{\infty}$ ,  $\beta_j = -1$  или  $\beta_j = +1$ ;  $K^1 = K$ ,  $C^1 = C$ ,  $K^{-1} = \bar{K}$ ,  $C^{-1} = \bar{C}$ . Так как  $H$  не пусто, то существует такая пара  $(\alpha, \beta)$ , что множество

$$H_{\alpha\beta} = \left[ \bigcap_{i=1}^{\infty} K_i^{\alpha_i} \right] \cap \left[ \bigcap_{j=1}^{\infty} C_j^{\beta_j} \right]$$

также не пусто и, следовательно, содержит некоторый элемент  $x_0(t)$ . Представим теперь  $H_{\alpha\beta}$  в виде

$$H_{\alpha\beta} = \left[ \bigcap_{j: b_j=1} C_j^{\beta_j} \right] \cap \left[ \bigcap_{j: b_j=-1} \bar{C}_j^{\beta_j} \right] \cap \left[ \bigcap_i K_i^{\alpha_i} \right].$$

Каждое  $K_i$  накладывает ограничение только на  $x(t_i)$ . Для тех номеров  $j$ , у которых  $\beta_j = -1$ , выберем такое  $\tau_j \in T_j$ , чтобы  $x_0(\tau_j) \notin F_j$ . Объединим обе последовательности  $\{t_i\}$  и  $\{\tau_j\}$  в одну последовательность  $\{t_k\}$ . Для каждого  $C_j$ , у которого  $\beta_j = +1$ , введем функцию  $f_j(t)$ :  $f_j(t) = -\infty$  при  $t \notin T_j$ ,  $f_j(t) = \min x \in F_j$  при  $t \in T_j$ .

Ступенчатая функция  $f_j(t)$  является, очевидно, функцией первого класса по Бэру. Обозначим  $f(t) = \sup_j f_j(t)$  и положим  $y(t) = f(t)$  при  $t \notin \{t_k\}$ ,  $y(t) = x_0(t)$  при  $t = t_k$ . Очевидно, что  $y(t) \in K_i^{\alpha_i}$ , так как это однозначно определяется по значению функции при  $t = t_i$ , а  $y(t_i) = x_0(t_i)$ . Далее,  $y(t) \in \bar{C}_j$  при  $\beta_j = -1$ , так как  $y(\tau_j) \notin F_j$ , и, следовательно,  $y(t)$  не лежит всюду внутри коридора  $C_j$ . Наконец,  $f(t) \in C_j$  при всех  $j$  с  $\beta_j = +1$  по построению, так что  $y(t)$  при  $t \notin \{t_k\}$  лежит внутри коридора  $C_j$ . А при  $t = t_k$   $y(t_k)$  совпадает с  $x_0(t_k)$  и тоже лежит внутри  $C_j$ . Поэтому  $y(t) \in H_{\alpha\beta}$ . Так как  $y(t)$  отличается от функции  $f(t)$  второго класса лишь на счетном множестве значений, то  $y(t)$  также является функцией второго класса по Бэру.

Представляется интересным выяснить, какие функциональные множества являются дубовскими, а какие нет. Можно показать, что являются дубовскими:

- 1) множество функций, непрерывных в данной точке (или непрерывных только слева или только справа);
- 2) множество, состоящее из одной функции без разрывов второго рода;
- 3) множество всех непрерывных функций;
- 4) множество всех функций с разрывами только первого рода, принимающих в точке разрыва одно из предельных значений.

Не являются дубовскими:

- 5) любое множество  $A$ ,  $A \neq R^T$ , содержащее все бэровские функции нулевого, первого и второго классов (в частности, множество всех бэровских функций, множество всех функций, измеримых по Лебегу, и т.п.);
- 6) множество, состоящее из одной функции  $x(t)$ , не являющейся бэровской функцией нулевого, первого или второго классов.

Отрицательные утверждения 5) и 6) легко вывести, применив теорему 4 к множествам  $\bar{A}$  и к  $C = x(t)$ . Утверждение 1) тривиально. Чтобы доказать 2), достаточно заметить, что всякая функция без разрывов второго рода является монотонным пределом (как сверху, так и снизу) последовательности ступенчатых функций. Доказательство утверждений 3) и 4) сложнее и основывается на следующих характеристических свойствах указанных семейств функций:

чтобы функция  $x(t)$ ,  $0 \leq t \leq 1$ , была непрерывной на  $[0, 1]$ , необходимо и достаточно, чтобы ее модуль непрерывности  $\varphi(h) \rightarrow 0$  при  $h \rightarrow 0$ ;

чтобы функция  $x(t)$ ,  $0 \leq t \leq 1$ , не имела разрывов второго рода на  $[0, 1]$ , необходимо и достаточно, чтобы ее модуль квазинепрерывности  $\psi(h) \rightarrow 0$  при  $h \rightarrow 0$  (см. [6]).

Проводить соответствующие конструкции мы не будем.

## СПИСОК ЛИТЕРАТУРЫ

1. Колмогоров А.Н. Основные понятия теории вероятностей. — М.-Л.: ОНТИ, 1936.
2. Ченцов Н.Н. Топологические меры и теория случайных функций // Труды Всесоюзного совещания по теории вероятности. Ереван, 1960. — С. 83–87.

3. Халмош П. Теория меры. — М.: ИЛ, 1953.
4. Дуб Дж.Л. Вероятностные процессы. — М.: ИЛ, 1956.
5. Матн Н.В. On the realization of stochastic processes by probability distribution in function spaces // Sankhyā. 1951. V. 11, № 1. P. 3–8.
6. Ченцов Н.Н. Предельные теоремы для некоторых классов случайных функций // Труды Всесоюзного совещания по теории вероятностей. Ереван, 1960. — С. 280–285.

## СЛУЧАЙНЫЕ ПОЛЯ ЛЕВИ<sup>\*</sup>)

Рассмотрим на метрическом пространстве  $G$  сепарабельное случайное поле со следующими свойствами.

1. Совместное распределение значений поля  $\xi(A_1), \dots, \xi(A_n)$  является нормальным для каждого конечного набора точек  $A_1, A_2, \dots, A_n$ , принадлежащих  $G$ .
2.  $\mathbf{M}\xi(A) = 0$ .
3.  $\mathbf{M}[\xi(A_i) - \xi(A_j)]^2 = \rho(A_i, A_j)$ , где  $\rho(A_i, A_j)$  — расстояние между точками  $A_i$  и  $A_j$ .

Условия 1–3 определяют случайное поле с точностью до нормально распределенного случайного слагаемого. Для определенности можно в некоторой фиксированной точке  $A_0$  положить  $\xi(A_0) = 0$ . Так определенное поле назовем *полем Леви*. Поле Леви является одним из возможных обобщений винеровского случайного процесса на случайные функции нечислового аргумента. Впервые такое обобщение было сделано П. Леви [1], который рассмотрел случайное поле с векторным аргументом под названием броуновского движения от нескольких параметров и изучил его свойства.

В заметке обсуждается вопрос, на каких поверхностях в целом существует случайное поле со свойствами 1–3.

Условия 1 и 2 выражают, что совместное распределение  $\xi(A_0) \equiv 0, \xi(A_1), \dots, \xi(A_n)$  является многомерным нормальным несмещенным распределением с матрицей вторых моментов  $\|\mathbf{M}[\xi(A_i)\xi(A_j)]\|_{i,j=0}^n$  где, как нетрудно вычислить,

$$\mathbf{M}[\xi(A_i)\xi(A_j)] = \frac{1}{2}[\rho(A_0, A_i) + \rho(A_0, A_j) - \rho(A_i, A_j)].$$

Таким образом, для того чтобы на пространстве  $G$  существовало поле Леви, необходимо и достаточно, чтобы матрица вторых моментов была неотрицательно определенной при любом наборе точек  $A_1, \dots, A_n$ , т.е. чтобы

$$\sum_{i,j=0}^n [\rho(A_0, A_i) + \rho(A_0, A_j) - \rho(A_i, A_j)] z_i z_j \geq 0$$

<sup>\*</sup>) Теория вероятн. и ее примен. 1968. Т. 13, вып. 1. С. 152–155 (совм. с Е.А. Морозовой).

при любых  $n; A_1, \dots, A_n; z_1, \dots, z_n$  (при этом нулевое подпространство соответствует линейным детерминированным зависимостям). Непосредственно проверить это условие даже для евклидовой плоскости, как это сделано в [2], достаточно сложно.

Простая геометрическая конструкция случайного поля на плоскости была предложена Н.Н. Ченцовым [3] и затем усовершенствована и перенесена на поля на сфере П. Леви [4]. Там же был поставлен вопрос: на каких поверхностях и многообразиях существуют случайные поля со свойствами 1–3?

Изложим сначала усовершенствованную конструкцию поля Леви на плоскости.

Пусть  $F$  — произвольное множество,  $\sigma$  — некоторая алгебра его подмножеств,  $\mu$  — мера на измеримом пространстве  $(F, \sigma)$ . Рассмотрим белый шум, порожденный мерой  $\mu$ , т.е. случайную “меру”  $\eta(B)$  на  $(F, \sigma)$  задаваемую условиями (ср. [5]):

$$0) \eta(B_1 \cup B_2) = \eta(B_1) + \eta(B_2) - \eta(B_1 \cap B_2);$$

1а) совместное распределение  $\eta(B_1), \dots, \eta(B_n)$  является нормальным при всех  $n, B_1, \dots, B_n$ ;

$$2а) \mathbf{M}\eta(B) = 0;$$

$$3а) \mathbf{M}[\eta(B_1)\eta(B_2)] = \mu(B_1 \cap B_2).$$

Существование нормальной случайной функции очевидно, так как для непересекающихся множеств  $B_1, \dots, B_n$  матрица вторых моментов диагональна, а в условии 3а) обе части равенства дистрибутивны относительно объединения непересекающихся множеств.

В курсах интегральной геометрии (см., например, [6]) доказывается, что во множестве ориентированных прямых евклидовой плоскости можно ввести меру, инвариантную относительно движений, так, чтобы мера множества прямых, пересекающих отрезок  $A_1A_2$ , равнялась длине отрезка.

Как обычно, пару направлений  $l, t$  называем *правой парой*, если поворот на наименьший угол от  $l$  до  $t$  производится против часовой стрелки, *левой*, если направление вращения по часовой стрелке. Обозначим  $\{A_0A\}^+$  и  $\{A_0A\}^-$  — множества ориентированных прямых, пересекающих ориентированный отрезок  $A_0A$  и образующих с ним правую и соответственно левую пару. Рассмотрим отвечающий инвариантной мере белый шум  $\eta(\cdot)$  на множестве ориентированных прямых и выберем на плоскости точку  $A_0$  за начальную.

**ТЕОРЕМА.** Поле  $\xi(A) = \eta\{A_0A\}^+ - \eta\{A_0A\}^-$  является полем Леви.

**Доказательство.** Условие 1) следует из 1а), условие 2) — из 2а). Остается проверить 3). Из аксиомы Паша следует, что если прямая пересекает одну сторону треугольника, то она пересекает еще одну его сторону. И поэтому множество прямых, пересекающих треугольник  $A_0A_1A_2$ , разбивается с точностью до меры ноль на три непересекающихся класса прямых

$$\{A_0A_1 \cap A_1A_2\}, \quad \{A_0A_1 \cap A_0A_2\}, \quad \{A_0A_2 \cap A_1A_2\}.$$

Мера прямых, проходящих через вершину, равна нулю, причем

$$\begin{aligned} \{A_iA_j\} \cap \{A_jA_k\} &= (\{A_iA_j\}^+ \cap \{A_jA_k\}^-) \cup (\{A_iA_j\}^- \cap \{A_jA_k\}^+), \\ (\{A_iA_j\}^\mp \cap \{A_jA_k\}^\pm) \cup (\{A_jA_k\}^\pm \cap \{A_kA_i\}^\pm) &= \{A_jA_k\}^\pm. \end{aligned}$$

Вычислим  $\xi(A_1) - \xi(A_2)$ . Имеем

$$\begin{aligned}
\xi(A_1) - \xi(A_2) &= \eta\{A_0A_1\}^+ - \eta\{A_0A_1\}^- - \eta\{A_0A_2\}^+ + \eta\{A_0A_2\}^- = \\
&= \eta(\{A_0A_1\}^+ \cap \{A_1A_2\}^-) + \eta(\{A_0A_1\}^+ \cap \{A_0A_2\}^+) - \\
&\quad - \eta(\{A_0A_1\}^- \cap \{A_1A_2\}^+) - \eta(\{A_0A_1\}^- \cap \{A_0A_2\}^+) - \\
&\quad - \eta(\{A_0A_2\}^+ \cap \{A_2A_1\}^-) - \eta(\{A_0A_2\}^+ \cap \{A_0A_1\}^+) + \\
&\quad\quad + \eta(\{A_0A_2\}^- \cap \{A_2A_1\}^+) + \eta(\{A_0A_2\}^- \cap \{A_0A_1\}^-) = \\
&= \eta(\{A_0A_1\}^+ \cap \{A_1A_2\}^-) + \eta(\{A_0A_2\}^- \cap \{A_2A_1\}^+) - \\
&\quad - \eta(\{A_0A_1\}^- \cap \{A_1A_2\}^+) + \eta(\{A_0A_2\}^+ \cap \{A_2A_1\}^-) = \\
&\quad\quad = \eta\{A_1A_2\}^- - \eta\{A_1A_2\}^+,
\end{aligned}$$

откуда  $\mathbf{M}[\xi(A_1) - \xi(A_2)]^2 = \rho(A_1, A_2)$ .

Указанная конструкция поля Леви автоматически распространяется на достаточно малые геодезически выпуклые области гладкой поверхности. Если только в такой окрестности через любые две точки проходит ровно одна геодезическая и из каждой точки на геодезическую опускается ровно один перпендикуляр, то можно так ввести плотность множества геодезических, что мера множества геодезических, пересекающих данный отрезок геодезической, будет равняться его длине [6].

На односвязных поверхностях неположительной кривизны, как следует из теоремы Гаусса–Бонне, оба эти свойства выполнены в целом. Поэтому на односвязных поверхностях неположительной кривизны естественная плотность геодезических существует в целом, и, следовательно, существует в целом поле Леви. На поверхности с сопряженными точками эта конструкция очевидным образом не распространяется. Более того, ниже будут приведены примеры поверхностей, на которых поля Леви не существуют.

Приведем необходимое условие существования поля Леви на классе поверхностей, у которых имеется замкнутая геодезическая  $S$  такая, что любая ее половина является кратчайшей. Очевидно, что любая пара ее диаметрально противоположных точек будет сопряженной. Для таких поверхностей имеет место

**ТЕОРЕМА.** Пусть на поверхности  $G$  существует замкнутая геодезическая  $S$  такая, что любая ее половина является кратчайшей. Для того чтобы на  $G$  существовало поле Леви, необходимо, чтобы для каждой точки  $A \in G$  сумма расстояний до любой пары диаметрально противоположных точек  $A_1, A_2 \in S$  была постоянна, т.е. чтобы

$$\rho(A, A_1) + \rho(A, A_2) \equiv c(A).$$

**Доказательство.** Пусть поле Леви  $\xi(A)$  существует; тогда оно вдоль спрямляемой кривой  $S$ , как следует из критерия Колмогорова [7], непрерывно. Поэтому оно с вероятностью 1 интегрируемо на  $S$ , и мы можем, положив  $\xi_1(A) = \xi(A) - \frac{1}{2r} \oint_S \xi(A(s)) ds$ , где  $2r$  — длина  $S$ , рассматривать другое поле, для которого  $\oint_S \xi(A(s)) ds = 0$ . Это поле удовлетворяет условиям 1)–3) и поэтому является полем Леви.

Рассмотрим такое поле  $\xi(A)$  на  $S$ . Выберем за начало отсчета дуги произвольную точку  $O$ . Вычисляем:

$$\begin{aligned} \mathbf{M}\xi^2(O) &= \mathbf{M}\left[\frac{1}{2r} \int_{-r}^{+r} (\xi(O) - \xi(s)) ds\right]^2 = \\ &= \frac{1}{4r^2} \int_{-r}^{+r} \int_{-r}^{+r} \mathbf{M}[\xi(O) - \xi(s)][\xi(O) - \xi(t)] ds dt = \\ &= \frac{1}{4r^2} \int_{-r}^{+r} \int_{-r}^{+r} \frac{1}{2} \{-\mathbf{M}[\xi(t) - \xi(s)]^2 + \mathbf{M}[\xi(O) - \xi(s)]^2 + \mathbf{M}[\xi(O) - \xi(t)]^2\} ds dt = \\ &= \frac{1}{8r^2} \int_{-r}^{+r} 2 \int_0^{+r} s ds dt = \frac{1}{8r^2} \int_{-r}^{+r} r^2 dt = \frac{1}{4} r. \end{aligned}$$

Таким образом,  $\mathbf{M}\xi^2(O) = r/4$  независимо от положения точки на геодезической. Пусть теперь  $A_1, A_2$  — произвольная пара диаметрально противоположных точек. Тогда

$$\begin{aligned} \mathbf{M}[\xi(A_1) + \xi(A_2)]^2 &= 2\mathbf{M}[\xi(A_1)]^2 + 2\mathbf{M}[\xi(A_2)]^2 - \mathbf{M}[\xi(A_1) - \xi(A_2)]^2 = \\ &= 4\mathbf{M}\xi^2(O) - r = 0. \end{aligned}$$

Таким образом,  $\xi(A_1) + \xi(A_2) = 0$  с вероятностью 1. Пусть  $A$  — произвольная точка  $G$ . Тогда

$$\begin{aligned} 4\mathbf{M}[\xi(A)]^2 &= \mathbf{M}[2\xi(A) - \xi(A_1) - \xi(A_2)]^2 = \\ &= \mathbf{M}\{\xi(A) - \xi(A_1) + \xi(A) - \xi(A_2)\}^2 + 4\mathbf{M}[\xi(A_1)]^2 - r = \\ &= \mathbf{M}\{\xi(A) - \xi(A_1) + [\xi(A) - \xi(A_2)]\}^2 + \mathbf{M}\{[\xi(A) - \xi(A_1)] - [\xi(A) - \xi(A_2)]\}^2 - r = \\ &= 2\mathbf{M}[\xi(A) - \xi(A_1)]^2 + 2\mathbf{M}[\xi(A) - \xi(A_2)]^2 - r = 2\rho(A, A_1) + 2\rho(A, A_2) - r. \end{aligned}$$

Следовательно,

$$\rho(A, A_1) + \rho(A, A_2) = \frac{r}{2} + 2\mathbf{M}[\xi(A)]^2 \equiv c(A).$$

Как легко видеть, это условие выполнено для сферы с  $c(A) \equiv r$  и не удовлетворяется ни для цилиндра, ни для тора, ни для эллипсоида, отличного от сферы.

#### СПИСОК ЛИТЕРАТУРЫ

1. Lévy P. Le mouvement brownien plan // Amer. J. Math. 1940. V. 62. P. 487–550.
2. Lévy P. Processus stochastiques et mouvement brownien. — Paris: Gauthier-Villars, 1948.
3. Ченцов Н.Н. Многопараметрическое броуновское движение Леви и обобщенный белый шум // Теория вероят. и ее примен. 1957. Т. 2, вып. 2. С. 281–282.
4. Lévy P. Le mouvement brownien fonction d'un point de la sphère de Riemann // Rend. Circolo mat. de Palermo. 1959. V. 8, № 3. P. 297–310.
5. Prékopa A. On stochastic set functions, I // Acta Math. Acad. Sci. Hung. 1956. V. VII, № 2. P. 215–263.

6. Сантало Л.А. Введение в интегральную геометрию. — М.: ИЛ, 1956.  
 7. Слуцкий Е.Е. Qualche propositione relativa alla teoria delle funzioni aleatorie // Giornale dell' Istituto Italiano degli Attuari. 1937. V. 8, № 2. P. 183–199.

## НЕИЗМЕРИМЫЕ ПОДМНОЖЕСТВА ТИХОНОВСКОГО КУБА<sup>\*)</sup>

Понятие тихоновского произведения системы топологических пространств [1] является одним из основных понятий общей топологии и функционального анализа. В частности, пространство всех функций аргумента  $t \in T$  со значениями в топологическом пространстве  $Y$  естественно рассматривать как топологизированную по Тихонову степень  $Y^T$ . Вероятностная мера, заданная на некоторой  $\sigma$ -алгебре подмножеств тихоновской степени  $Y^T$ , определяет случайные  $Y$ -функции аргумента  $t$ . Поэтому теория случайных процессов и смежные разделы теории вероятностей и математической статистики существенно опираются на теорию измеримых множеств в топологических пространствах  $Y^T$ . В этой работе мы ограничимся пространствами  $\mathfrak{R}^T$  вещественнозначных функций с несчетным множеством  $T$  индексов  $t$  и их подпространствами  $E^T$ , где  $E$  — единственный отрезок из  $\mathfrak{R}$  с обычной топологией; последние принято называть [2] *тихоновскими кубами*.

Согласно [1], множества вида  $\mathfrak{R}^{T-t} \times I_t$  (соответственно  $E^{T-t} \times I_t$ ) при всевозможных  $t \in T$  и интервалах  $I = I_t$  образуют открытую предбазу пространства  $\mathfrak{R}^T$ , а всевозможные конечные пересечения таких множеств — открытую базу. На вещественной прямой  $\mathfrak{R}$  любой интервал и любое открытое множество являются множеством типа  $F_\sigma$ , т.е. не более чем счетными объединениями замкнутых множеств. Для несчетных степеней дело обстоит сложнее. Базовые открытые множества по-прежнему имеют также тип  $F_\sigma$ , но для любых открытых множеств это уже неверно. Поэтому в пространствах  $\mathfrak{R}^T$  приходится различать  $\sigma$ -алгебру  $\mathfrak{B}_0$  бэровских множеств, порожденную классом базовых открытых множеств, и  $\sigma$ -алгебру  $\mathfrak{B}$  борелевских множеств, порожденную классом всех открытых множеств [3].

Согласно фундаментальной теореме Колмогорова [4], всякая  $\sigma$ -аддитивная конечная мера, заданная на порожденной базовыми открытыми множествами алгебре  $\mathfrak{B}'_0$ , всегда продолжается, и притом единственным образом, на всю  $\sigma$ -алгебру  $\mathfrak{B}_0$ . Таким образом, семейство согласованных конечномерных распределений вероятностей однозначно определяет случайный процесс и позволяет вычислять математические ожидания функционалов от траектории процесса, непрерывных в тихоновской топологии.

<sup>\*)</sup> Проблемы математической физики и вычислительной математики. — М.: Наука, 1977. — С. 316–320.

Всякое бэровское множество  $A$  обязательно представимо в виде  $\mathfrak{R}^{T-\tau} \times A_0$ , где множество  $\tau$  индексов не более чем счетно, а  $A_0$  — борелевское подмножество  $\mathfrak{R}^\tau$ . Так как тихоновские кубы  $F^T$ ,  $F = [x: a \leq x \leq b]$ , не удовлетворяют этому описанию, то множество “попадающих в коридор”

$$\{\xi(\cdot): a \leq \xi(t) \leq b \quad \forall t \in T\}$$

траекторий процесса неизмеримо  $\mathfrak{B}_0$ . Но оно измеримо по Борелю, ибо по знаменитой<sup>1)</sup> теореме А.Н. Тихонова множество компактно и тем самым замкнуто.

По теореме Какутаки [5] (см. также [3]) у всякой бэровской меры на  $Y^T$ , где  $Y$  — компакт со счетной базой, существует главное продолжение — продолжение до регулярной борелевской меры. На компактах она совпадает с внешней бэровской мерой, а на любом открытом множестве равна верхней грани мер всевозможных счетных объединений содержащихся в нем базовых. (Это свойство мы предложили в [6] называть *несчетной аддитивностью* топологической меры.) В подавляющем большинстве приложений интересно именно это регулярное продолжение. Для пространства  $\mathfrak{R}^T$  аналогичное утверждение несправедливо, так что траектории неограниченного регулярного случайного процесса приходится трактовать как точки в  $[\mathfrak{R}^T]$ , где компактификация  $[\mathfrak{R}] = \{x: -\infty \leq x \leq +\infty\}$ . Поскольку  $[\mathfrak{R}]$  гомеоморфно  $E$ , для теории случайных процессов достаточно изучить неизмеримые подмножества куба  $E^T$ .

Класс  $\mathfrak{B}$ -измеримых подмножеств тихоновского куба  $E^T$  весьма широк. Когда  $T$  — вещественный интервал, то  $\mathfrak{B}$ -измеримы, например: множество функций, непрерывных в данной точке (или непрерывных только слева или только справа); множество всех непрерывных функций; множество, состоящее из одной функции; множество всех функций с разрывами только первого рода, принимающих в каждой точке разрыва одно из двух предельных значений [6–8]. Однако в отличие от случая отрезка  $E$  или конечномерного куба  $E^n$ , где неизмеримые множества неконструктивны, при несчетном  $T$  существуют (как было аннотировано нами в [6]) очень простые конструктивные примеры неизмеримых множеств. Таковы, например, все открытые ящики

$$(a, b)^T = \{x(\cdot): a < x_t < b \quad \forall t \in T\}.$$

Условимся о терминологии. Как известно [3], всякую вероятностную меру  $P$  на  $\sigma$ -алгебре  $S$  множеств можно продолжить по Лебегу до полной меры  $P^*$  на  $\sigma$ -алгебре  $S_P^* \supseteq S$ . Класс  $S_P^*$  зависит от меры  $P$ . Пересечение  $S^* = \bigcap_P S_P^*$  по всем вероятностным мерам  $P$  на  $S$  назовем *замыканием* или *абсолютным пополнением  $\sigma$ -алгебры* [9, § 2]. В замыкание  $\mathfrak{B}^*$  алгебры борелевских множеств входит также класс  $\mathfrak{A}$ -множеств и т.п.

**ТЕОРЕМА 1.** При  $\text{card } T \geq \mathfrak{c}$  открытый ящик  $(a, b)^T = \{x(\cdot): a < x_t < b \quad \forall t \in T\}$  неизмерим  $\mathfrak{B}^*$  в  $[\mathfrak{R}]^T$ .

**ТЕОРЕМА 2.** Если  $T = \{t: 0 \leq t \leq 1\}$ , то множество  $C_1$  всех функций без разрывов второго рода и непрерывных слева в каждой точке разрыва неизмеримо  $\mathfrak{B}^*$ .

<sup>1)</sup> “Классическая теорема А.Н. Тихонова о произведении бикompактных пространств, несомненно, является самой полезной теоремой о бикompактности ...” [10, с. 193].

Для доказательства достаточно указать в  $[\mathfrak{R}]^T$  две меры, совпадающие на всех борелевских множествах, но различающиеся на  $(a, b)^T$  для теоремы 1 и соответственно на  $C_l$  для теоремы 2. Перейдем к конструированию таких мер.

Пусть  $\Omega = \{\omega: 0 \leq \omega \leq 1\}$  — единичный отрезок,  $\mathfrak{L}$  — алгебра его лебеговых множеств, а  $\lambda\{\cdot\}$  — лебегова мера на  $(\Omega, \mathfrak{L})$ , т.е. “длина” множества. Для каждого отображения  $f: \Omega \rightarrow [\mathfrak{R}]^T$  возникает класс подмножеств  $A$  из  $[R]^T$ , имеющих  $\mathfrak{L}$ -измеримый прообраз  $f^{-1}(A)$ . Как известно [3], этот класс является  $\sigma$ -алгеброй и на нем индуцируется мера  $\lambda f^{-1}$ . Для наших целей нужны такие отображения  $f$ , чтобы прообразы  $\mathfrak{B}$ -множеств были измеримы  $\mathfrak{L}$ . Для этого достаточно, чтобы были  $\mathfrak{L}$ -измеримы прообразы всех открытых подмножеств тихоновского куба  $[\mathfrak{R}]^T$ .

Для построения близких отображений нам потребуется вспомогательное предложение, принадлежащее Морсу [10]. Условимся называть *собственным интервалом* числовой прямой полуоткрытый, замкнутый или открытый интервал, содержащий более одной точки.

**ЛЕММА 1.** *Объединение  $\bigcup_{\alpha} D_{\alpha}$  любого количества собственных интервалов  $D_{\alpha}$  отличается от объединения  $\bigcup_{\alpha} \langle D_{\alpha} \rangle$  их открытых ядер  $\langle D_{\alpha} \rangle$  не более чем на счетное множество точек.*

**СЛЕДСТВИЕ.**  $\lambda\{\bigcup_{\alpha} D_{\alpha}\} = \lambda\{\bigcup_{\alpha} \langle D_{\alpha} \rangle\}$ .

**ДОПОЛНЕНИЕ.**  $\bigcup_{\alpha} D_{\alpha} = \bigcup_i D_{\alpha(i)}$  для некоторой не более чем счетной последовательности индексов.

Для доказательства леммы заметим, что  $\bigcup_{\alpha} \langle D_{\alpha} \rangle$  открыто и потому разлагается в не более чем счетное объединение попарно непересекающихся интервалов  $I_n$ . При этом концевая точка  $D_{\alpha}$  оказывается либо внутренней точкой для  $I_n \supseteq \langle D_{\alpha} \rangle$ , либо концом этого  $I_n$ . Так как каждый интервал  $I_n$  есть объединение счетного числа компактов, то для его покрытия требуется не более чем счетное число множеств  $\langle D_{\alpha} \rangle$ . Объединяя их для всех  $n$ , получаем последовательность  $\alpha(i)$ .

По смыслу  $[\mathfrak{R}]^T$  отображение  $f: \Omega \rightarrow [\mathfrak{R}]^T$  ставит каждому  $\omega \in \Omega$  набор координат  $x_t(\omega) \forall t \in T$ , т.е. описывается числовой функцией  $x(t, \omega)$  на  $T \times \Omega$ . Выберем следующие две функции:

$$x_1(t, \omega) = 0 \quad \forall t < \omega, \quad x_1(t, \omega) = 1 \quad \forall t \geq \omega, \quad (1)$$

$$x_2(t, \omega) = 0 \quad \forall t \leq \omega, \quad x_2(t, \omega) = 1 \quad \forall t > \omega, \quad (2)$$

задающие отображения  $f_1$  и  $f_2$ .

**ЛЕММА 2.** *Прообраз любого предбазового открытого множества  $G = [\mathfrak{R}]^{T-t} \times I_t$  может быть пустым либо собственным промежутком, либо сводиться к точке  $\omega = 0$  или к точке  $\omega = 1$ . Прообразы  $f_1^{-1}(G)$  и  $f_2^{-1}(G)$  могут различаться не более чем на точку (концевую,  $\omega = 0$ ,  $\omega = 1$ ).*

**СЛЕДСТВИЕ.** *Индукцированные меры  $\lambda f_1^{-1}$  и  $\lambda f_2^{-1}$  совпадают на  $\mathfrak{B}$ -множествах. Они регулярны.*

**Доказательство.** Пусть  $I_t = (a, b)$ . Условие  $a < x(t, \omega) < b$  равносильно для  $x_1$  и  $x_2$  соответственно:

$$\begin{array}{lll}
0 \leq \omega \leq 1, & 0 \leq \omega \leq 1 & \text{при } a < 0 < 1 < b; \\
0 \leq \omega < t, & 0 \leq \omega \leq t & \text{при } a < 0 < b \leq 1; \\
t \leq \omega \leq 1, & t < \omega \leq 1 & \text{при } 0 \leq a < 1 < b; \\
\emptyset, & \emptyset & \text{при } 0 \notin (a, b), 1 \notin (a, b).
\end{array}$$

Таким образом, утверждения леммы для предбазовых открытых множеств выполнены. При этом для  $f_1^{-1}(G)$  левый конец промежутка обязательно входит в его состав, а правый, только когда он совпадает с  $\omega = 1$ . Для  $f_2^{-1}(G)$ , наоборот, левый конец входит только тогда, когда он есть  $\omega = 0$ . Пересечение конечного числа таких полузамкнутых промежутков одного типа — снова промежуток указанного типа. Он может сводиться к точке лишь при  $\omega = 0$  или  $\omega = 1$ . При этом для  $f_1$  и  $f_2$  прообразы отличаются максимум на концевые точки.

По определению тихоновского произведения любое открытое множество есть объединение некоторой системы базовых. Его  $f$ -прообраз есть объединение соответствующих промежутков. По лемме 1  $f_1$ -прообраз отличается от  $f_2$ -прообраза лишь на счетное число точек, а лебеговы меры прообразов совпадают. Регулярность (несчетная аддитивность) следует из дополнения. Так как открытые множества порождают  $\sigma$ -алгебру  $\mathfrak{B}$ , то индуцированные меры  $\lambda f_1^{-1}$  и  $\lambda f_2^{-1}$  совпадают на  $\mathfrak{B}$ . Лемма доказана.

**ТЕОРЕМА 3.** *Множество  $M_l$  монотонно неубывающих непрерывных слева числовых функций неотделимо  $\mathfrak{B}^*$  от множества  $M_r$  монотонно неубывающих непрерывных справа функций.*

**СЛЕДСТВИЕ.** *Множество  $M_l$  неизмеримо  $\mathfrak{B}^*$ .*

**Доказательство.** Раз индуцированные меры  $\lambda f_1^{-1}$  и  $\lambda f_2^{-1}$  совпадают на  $\mathfrak{B}$ , они совпадают и на  $\mathfrak{B}^*$ . Пусть  $\mathfrak{B}^* \ni A_l \supseteq M_l$ ,  $\mathfrak{B}^* \ni A_r \supseteq M_r$ ,  $A_l \cap A_r = \emptyset$ . Тогда  $\lambda f_i^{-1}(M_l) \leq \lambda f^{-1}(A_l)$ ,  $\lambda f_i^{-1}(M_r) \leq \lambda f^{-1}(A_r)$ ,  $i = 1, 2$ , где мера  $A$  не зависит от  $i$ . При этом должно быть  $\lambda f^{-1}(A_l) + \lambda f^{-1}(A_r) \leq 1$ . Но поскольку  $f_2^{-1}(M_l) = \Omega$ ,  $f_1^{-1}(M_r) = \Omega$ , то  $\lambda f^{-1}(A_l) = 1$ ,  $\lambda f^{-1}(A_r) = 1$ . Полученное противоречие доказывает неотделимость. Множество  $M_l$  при этом не может быть  $\mathfrak{B}^*$ -измеримым, иначе оно само и его дополнение отделяли бы  $M_l$  и  $M_r$ .

**Доказательство теоремы 2.** Если бы множество  $C_l$  было измеримо  $\mathfrak{B}^*$ , оно бы отделяло  $M_l$  от  $M_r$ , чего нет.

**Доказательство теоремы 3** проводится сходным образом. Например, для проверки  $\mathfrak{B}^*$ -неизмеримости ящика  $(-1, +1)^{[0,1]}$  можно построить функции:

$$\begin{array}{ll}
x_1(t, \omega) = t - \omega + 1 & \forall t < \omega, & x_1(t, \omega) = t - \omega & \forall t \geq \omega, \\
x_2(t, \omega) = t - \omega + 1 & \forall t \leq \omega, & x_2(t, \omega) = t - \omega & \forall t > \omega.
\end{array}$$

Разумеется, доказанные теоремы легко обобщаются на более сложные фазовые пространства и более сложные пространства индексов. Нам казалось, однако, более привлекательным, следуя [1], провести рассуждения на легко обозримом объекте, тем более, что он универсален (см. [1]).

Несмотря на свою простоту, построенные примеры приводят к далеко идущим следствиям отрицательного характера. Прежде всего, они показывают, что значение координаты разрывного случайного процесса в момент разрыва не яв-

ляется борелевской величиной. Поэтому для вероятностного анализа таких величин стандартные подходы не годятся.

Приведенные примеры показывают также неприменимость прямого “борелевского” подхода в теории случайных мер. Совершенно естественно задавать случайную меру на прямой ее функцией распределения, т.е. случайной монотонно неубывающей функцией, непрерывной слева. Мы показали здесь, что множество таких функций является неизмеримым  $\mathfrak{B}^*$  (хотя множество непрерывных функций распределения  $\mathfrak{B}$ -измеримо). Поэтому для построения теории приходится прибегать к более сложным, “неборелевским” конструкциям [9, § 29] в тихоновском кубе  $E^S$ , где  $\sigma$ -алгебра измеримых множеств  $S$  — область задания случайной меры.

### СПИСОК ЛИТЕРАТУРЫ

1. *Tichonov A.N.* Über die topologische Erweiterung von Räumen // *Math. Ann.* 1929. V. 102. P. 544–561.
2. *Марков А.А.* Топология // *Математика в СССР за тридцать лет.* — М.-Л.: Гостехиздат, 1948.
3. *Halmos P.R.* *Measure theory.* — N.Y., 1950.
4. *Kolmogorov A.N.* *Grundbegriffe der Wahrscheinlichkeitsrechnung.* — Berlin, 1933.
5. *Kakutani S.* Notes on infinite product measure spaces. V. I // *Proc. Imper. Acad.* 1943. V. 19. P. 148–151.
6. *Ченцов Н.Н.* Топологические меры и теория случайных функций // *Труды V Всесоюз. совещ. по теории вероятн. и мат. статистике (Ереван, 1958).* — Ереван: Изд-во АН АрмССР, 1960. — С. 83–87.
7. *Ченцов Н.Н.* Дубовские множества и дубовские распределения вероятностей // *Труды VI Всесоюз. совещ. по теории вероятн. и мат. статистике (Вильнюс, 1960).* — Вильнюс: Госиздат ЛитССР, 1962.
8. *Meyer A.* Separabilité d'un processus stochastique // *C. r. Acad. Sci.* 1959. V. 249. P. 2475–2476.
9. *Ченцов Н.Н.* Статистические решающие правила и оптимальные выводы. — М.: Наука, 1972.
10. *Kelley J.L.* *General topology.* — N.Y., 1957.

РАЗДЕЛ II

**ТЕОРИЯ КУБАТУРНЫХ  
ФОРМУЛ.  
МЕТОД МОНТЕ-КАРЛО**

---

**О ЧИСЛЕННОМ ВЫЧИСЛЕНИИ  
КОНТИНУАЛЬНЫХ ИНТЕГРАЛОВ<sup>\*)</sup>**

Известно, что ответы многих физических задач можно записать в виде так называемых континуальных интегралов — интегралов в функциональных пространствах. Однако самое вычисление континуального интеграла вызывает ряд трудностей, и только в редких случаях можно довести аналитические выкладки до числа. Поэтому неоднократно высказывалось мнение о практической бесполезности такого ответа.

Следует, однако, ожидать, что, как только будут разработаны приемы численного континуального интегрирования, решение задач при помощи континуальных интегралов приобретет практический интерес. В настоящей заметке предлагается следующий путь расчетов. Континуальный интеграл, записываемый как интеграл по некоторой мере, аппроксимируется конечномерным интегралом Стильтьеса достаточно высокой кратности, который затем и вычисляется обычным образом. Этот путь оказывается вполне приемлемым, так как вычисление кратных интегралов при помощи современной вычислительной техники особых затруднений не вызывает.

В качестве примера такого расчета в настоящей заметке определяется величина  $E_0$  низшего энергетического уровня полярона — медленного электрона в ионном кристалле. Теорией поляронов много занимался Пекар [1], который, в частности, вычислил величину  $E_0$  в адиабатическом приближении. Фейнман показал [2], что определение  $E_0$  можно свести к вычислению асимптотики континуального интеграла

$$K = \int \exp SD\mathbf{x}(t), \quad (1)$$

---

<sup>\*)</sup>ЖЭТФ. 1956. Т. 31, вып. 6 (12). С. 1106–1107 (совм. с И.М. Гельфандом).

где функционал  $S$  зависит от константы связи  $\alpha$ :

$$S = -\frac{1}{2} \int_0^T \left( \frac{d\mathbf{x}}{dt} \right)^2 dt + \frac{\alpha}{2\sqrt{2}} \int_0^T \int_0^T |\mathbf{x}(t) - \mathbf{x}(s)|^{-1} e^{-|t-s|} dt ds, \quad (2)$$

$$\alpha = \frac{1}{2} \left( \frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon} \right) \left( \frac{e^2}{\hbar\omega} \right) \left( \frac{2m\omega}{\hbar} \right)^{1/2}.$$

Здесь  $\varepsilon$  и  $\varepsilon_\infty$  — статическая и высокочастотная диэлектрические постоянные. Тогда при  $T \rightarrow \infty$ ,  $K(T) \sim \exp\{-E_0 T\}$ . Для численного определения величины  $E_0$  как функции  $\alpha$  Фейнман воспользовался вариационными методами. Полученные им приближенные разложения  $E_0$  как функции  $\alpha$  приведены в [2]. При больших  $\alpha$  хорошо действует адиабатическое приближение Пекара [1].

Для наших расчетов выражение (1) было преобразовано. Экспоненциальный множитель

$$\exp \left\{ -\frac{1}{2} \int_0^T \left( \frac{d\mathbf{x}}{dt} \right)^2 dt \right.$$

был внесен под знак дифференциала, и тогда интеграл записывается в виде среднего значения некоторого функционала по мере Винера:

$$K(T) = \int \exp \left\{ \frac{\alpha}{2\sqrt{2}} \int_0^T \int_0^T |\mathbf{x}(t) - \mathbf{x}(s)|^{-1} e^{-|t-s|} dt ds \right\} d_W x(t) =$$

$$= \mathfrak{M}_W \left[ \exp \left\{ \frac{\alpha}{2\sqrt{2}} \int_0^T \int_0^T |\mathbf{x}(t) - \mathbf{x}(s)|^{-1} e^{-|t-s|} dt ds \right\} \right], \quad (3)$$

$$E_0(\alpha) = - \lim_{T \rightarrow \infty} \frac{1}{T} \ln K(T). \quad (4)$$

Нетрудно видеть, что предел

$$- \lim_{T \rightarrow \infty} \frac{1}{T} \mathfrak{M}_W [\exp\{\alpha I\}], \quad (5)$$

где

$$I = \frac{1}{\sqrt{2}} \int_0^T d\eta \int_0^E \sqrt{\xi} e^{-\xi} |\mathbf{x}(t) - \mathbf{x}(s)|^{-1} d\sqrt{\xi} \quad (6)$$

( $\xi = t - s$ ,  $\eta = t + s$ ), отличается от  $E_0(\alpha)$  сколь угодно мало, лишь бы  $\Xi$  было достаточно велико. В то же время порядок остаточного члена в (5) меньше, чем в (4). Так как  $\sqrt{\xi} e^{-\xi}$  быстро убывает с ростом  $\xi$ , то при приближенных расчетах достаточно брать интеграл по  $\xi$  до  $\Xi \sim 5-7$ . Среднее значение функционала  $e^{\alpha I}$  вычислялось методом Монте-Карло. Были рассчитаны два варианта: с  $T = 4,5$  и  $T = 9,0$ . В обоих вариантах  $\Xi = 5,0$ , что дает ошибку в интеграле  $I$  около 0,3% в сторону преуменьшения. С шагом  $\tau = 0,05$  от  $t = 0$  до  $t = T + \Xi$  определялись случайные положения трехмерной брауновской траектории. Область интегрирования разбивалась на прямоугольники, двойной интеграл по каждому прямоугольнику вычислялся по простой разностной схеме, куда входило только значение подинтегральной функции в центре прямоугольника. После этого для набора значений  $\alpha$  вычислялись  $e^{\alpha I}$  и их средние арифметические по

всем  $N$  траекториям. Таким образом, вместо континуального интеграла были считаны методом Монте-Карло аппроксимирующие 190-кратные и 280-кратные интегралы. Результаты после введения статических поправок даны в табл. 1 и 2.

Таблица 1

$T$	4,5	4,5	9,0	9,0
$N$	300	600	300	400
$\alpha$	$-E_0(\alpha)$			
0,5	0,5051	0,5048	0,5047	0,5043
1,0	1,0208	1,0195	1,0195	1,0192
2,0	2,0880	2,0819	2,0839	2,0851
4,0	4,3900	4,3587	4,3469	4,3538
5,0	5,6271	5,5787	5,5205	5,5305

Таблица 2

$T$	4,5	4,5	9,0	9,0	Точное значение по [3]
$N$	300	600	300	400	
$-\varkappa_1$	1,0063	0,9999	0,9912	0,9940	1,0000
$-\varkappa_2$	0,0397	0,0370	0,0354	0,0370	0,0252

Для контроля вычислялось среднее значение случайной величины и ее старшие моменты, по которым были найдены коэффициенты разложения функции  $E_0(\alpha)$  в ряд  $E_0(\alpha) = \frac{\varkappa_1}{1!}\alpha + \frac{\varkappa_2}{2!}\alpha^2 + \dots$

В работе принимали участие А.С. Фролов и В.Ю. Крылов. Программировали задачу В.Н. Торопцева и Т.И. Фролова. Авторы выражают благодарность М.Р. Шура-Бура за ценное обсуждение вопросов программирования задачи.

#### СПИСОК ЛИТЕРАТУРЫ

1. Пекар С.И. Исследования по электронной теории кристаллов. — М.: Гостехиздат, 1951.
2. Feutchan R.P. // Pys. Rev. 1955. V. 97. P. 660
3. Нага Е. // Progr. Theor. Phys. 1954. V. 11. P. 449.

## О КВАДРАТУРНЫХ ФОРМУЛАХ ДЛЯ ФУНКЦИЙ БЕСКОНЕЧНО БОЛЬШОГО ЧИСЛА ПЕРЕМЕННЫХ \*)

В работе изучаются квадратурные формулы для некоторых простейших классов функций от счетной последовательности аргументов. Цель работы — выяснить на простейших моделях, какого типа должны быть ограничения на

\*) ЖВМ и МФ. 1961. Т. 1, № 3. С. 418–424.

класс функций, чтобы для него существовали приемлемые квадратурные формулы.

В качестве области изменения аргументов возьмем счетномерный единичный гиперкуб  $H$ . Точки в  $H$  задаются последовательностями

$$\mathbf{x} = \{x^{(1)}, x^{(2)}, \dots, x^{(n)}, \dots\},$$

где  $0 \leq x^{(n)} \leq 1$ ,  $n = 1, 2, \dots$ . Можно сказать, что  $H$  есть прямое произведение счетного числа отрезков  $[0, 1]$ . За меру в  $H$  мы возьмем прямое произведение лебеговых мер (т.е. длин) на ребрах. Интеграл по  $H$  от функции  $f(\mathbf{x})$  по мере мы будем писать как  $\int_H f(\mathbf{x}) d\mu$ . Для конечномерных функций, которые зависят лишь от нескольких первых аргументов, наш интеграл превращается в обыкновенный кратный интеграл

$$\int_H f(\mathbf{x}) d\mu = \underbrace{\int_0^1 \dots \int_0^1}_{n} f(\mathbf{x}) dx^{(1)} \dots dx^{(n)}.$$

В определении куба  $H$  и в определении меры все координаты входят равноправно. Если рассматривать классы  $K$  функций, которые равноправно зависят от своих координат, то для каждой индивидуальной функции класса  $K$  последовательность значений, получающихся по квадратурным формулам, с увеличением числа узлов может сходиться к значению интеграла. Однако эта сходимость не может быть равномерной по всему классу  $K$ . Дадим точные определения.

ОПРЕДЕЛЕНИЕ 1. Назовем класс  $K$  функций на  $H$  *классом типа T*, если для любой перестановки натурального ряда чисел

$$\begin{pmatrix} 1 & 2 & \dots & n & \dots \\ i_1 & i_2 & \dots & i_n & \dots \end{pmatrix}$$

вместе с каждой функцией  $f(x^{(1)}, x^{(2)}, \dots, x^{(n)}, \dots) \in K$  класс  $K$  содержит и функцию  $f(x^{(i_1)}, x^{(i_2)}, \dots, x^{(i_n)}, \dots)$ .

ОПРЕДЕЛЕНИЕ 2. *Квадратурной формулой по  $N$  точкам* называется функционал

$$\Phi_N[f] = \Phi_N(f(\mathbf{x}_1); \dots; f(\mathbf{x}_N)).$$

Он задается функцией  $\Phi_N(z_1; \dots; z_N)$  от  $N$  числовых аргументов и  $N$  точками — “узлами”  $\mathbf{x}_1, \dots, \mathbf{x}_N$ .

Важный частный случай, когда  $\Phi_N$  — линейная функция:

$$\Phi_N[f] = \sum_{i=1}^N \alpha_i f(\mathbf{x}_i).$$

В этом случае квадратурная формула задается указанием узлов и весов.

Мы покажем, что в классах типа T плохо квадратуемые функции встречаются даже среди функций от двух переменных. Нам потребуется для этого еще одно определение.

ОПРЕДЕЛЕНИЕ 3. Будем говорить, что функция  $f(y)$ ,  $-1 \leq y \leq +1$ , обладает П-свойствами, если

$$f(y) \xrightarrow{y \rightarrow 0} f(0) = 0, \quad (\Pi_1)$$

$$\int_{-1}^{+1} f(y) [1 - |y|] dy = I[f] > 0. \quad (\Pi_2)$$

Примерами таких функций могут служить функции  $f(y) = y^2$  и  $f(y) = \sin^2 2\pi y$ .

ТЕОРЕМА 1. Пусть функция  $\Phi_N(z_1; \dots; z_N)$ , задающая квадратурную формулу, непрерывна в нуле, так что

$$\overline{\Phi}_N(z_1; \dots; z_N) \xrightarrow{\max |z_i| \rightarrow 0} \overline{\Phi}_N(0; \dots; 0) = 0. \quad (\text{H})$$

Тогда среди функций  $f_{m,n}(\mathbf{x}) = f(x^{(m)} - x^{(n)})$  на кубе  $H$ , где  $f(y)$  обладает П-свойствами, найдется плохо квадратуемая функция, для которой формула  $\Phi_N[f]$  дает значение интеграла с относительной ошибкой более 50 %:

$$\Phi_N[f_{m,n}] \leq \frac{1}{2} \int_H f_{m,n}(\mathbf{x}) d\mu = \frac{1}{2} I > 0. \quad (1)$$

СЛЕДСТВИЕ 1.1. Если класс функций, имеющий тип Т, содержит хотя бы одну функцию вида  $f(x^{(1)} - x^{(2)})$ , где  $f(y)$  обладают П-свойствами, то квадратурные формулы, удовлетворяющие условию (H), не могут сходиться на  $K$  равномерно. В частности, последнее утверждение справедливо, если класс  $K$  содержит многочлены или тригонометрические полиномы.

Доказательство теоремы 1. По условию (H) выберем такое  $\rho$ , чтобы

$$\sup_{\max |z_i| \leq \rho} \Phi_N(z_1; \dots; z_N) \leq \frac{1}{2} I[f]. \quad (2)$$

Выберем теперь по условию ( $\Pi_1$ ) такое  $k$ , что

$$\sup_{|y| < 2^{-k}} |f(y)| \leq \rho. \quad (3)$$

Далее условимся говорить, что последовательность  $\{y_1; \dots; y_N\}$ ,  $0 \leq y \leq 1$ , принадлежит ящику  $A_K(u_1, \dots, u_N)$ ,  $1 \leq u_i \leq 2^k$ , если  $u_i - 1 \leq y_i \cdot 2^k \leq u_i$  при всех  $i$ .

Разнесем теперь по ящикам последовательности  $\{x_1^{(k)}, \dots, x_N^{(k)}\}$ ,  $k = 1, \dots, M$ , первых  $M = 2^{kN} + 1$  координат точек  $\mathbf{x}_i$  — узлов нашей квадратурной формулы. Таких последовательностей  $M$  штук, а ящиков всего  $2^{kN}$ , т.е. на единицу меньше. Поэтому хотя бы в одном ящике окажется по крайней мере две последовательности. Пусть ящик  $A(u_1, \dots, u_N)$  содержит последовательности  $m$ -х и  $n$ -х координат. Очевидно, при всех  $i$

$$|x_i^{(m)} - x_i^{(n)}| < 2^{-k}.$$

Учитывая (3), при всех  $i$  имеем

$$|f(x_i)| = |f(x_i^{(m)} - x_i^{(n)})| \leq \rho.$$

Из (2) получаем, что

$$\Phi_N[f_{m,n}] \leq \frac{1}{2} I[f].$$

Но

$$\int_H f(\mathbf{x}) d\mu = \int_0^1 \int_0^1 f(x^{(m)} - x^{(n)}) dx^{(m)} dx^{(n)} = \int_{-1}^{+1} f(y) [1 - |y|] dy = I[f] > 0.$$

Откуда

$$\int_H f_{m,n}(\mathbf{x}) d\mu - \Phi_N[f] \geq \frac{1}{2} I[f],$$

т.е. относительная ошибка более 50%.

Чтобы уяснить возможные разумные постановки задачи, рассмотрим винеровский случайный процесс. Пусть  $\eta(t)$  — траектория условного винеровского процесса,  $\eta(0) = \eta(1) = 0$ . Разложим  $\eta(t)$  в ряд Фурье по ортонормальной системе  $\sqrt{2} \sin k\pi t$ . Можно показать, что случайные коэффициенты  $\alpha_k$  распределены каждое по нормальному закону независимо друг от друга со средним 0 и дисперсией  $1/k^2\pi^2$ , т.е. со средним квадратичным уклонением  $1/k\pi$ . Таким образом, в смысле  $L_2$ -нормы в пространстве траекторий все координаты  $\alpha_k$  имеют одинаковый масштаб

$$\int_0^1 [\eta(t)]^2 dt = \sum_{k=1}^{\infty} \alpha_k^2,$$

а характерная длина относительно меры убывает как  $1/k$ ; плотность вероятности характеризуется квадратичной формой

$$\frac{1}{2} \sum_{k=1}^{\infty} k^2 \pi^2 \alpha_k^2,$$

так как плотность нормального распределения

$$f(\mathbf{a}) = \prod_k \left[ \frac{k^2 \pi^2}{2\pi} \right]^{1/2} \exp \left[ -\frac{1}{2} \sum_k k^2 \pi^2 \alpha_k^2 \right].$$

(Поверхность  $z = f(\mathbf{a})$  представляет собой гребень, склоны которого обрываются тем круче, чем больше номер координаты, по линии которой происходит спуск.)

Если сделать замену переменных  $z_k = k\pi\alpha_k$ , то квадратичная форма меры превратится в  $\frac{1}{2} \sum_{k=1}^{\infty} z_k^2$ , а квадратичная форма длины — в  $\sum_{k=1}^{\infty} z_k^2/k^2\pi^2$ .

Как известно, траектории винеровского процесса непрерывные, но не гладкие. То же самое можно сказать о траекториях условного винеровского процесса. Если  $\xi(t)$  — траектория винеровского процесса, то, как можно показать,  $\eta(t) = \xi(t) - t\xi(1)$ .

Далее возьмем

$$\Xi(t) = \int_0^1 \xi(\tau) d\tau, \quad \zeta(t) = \Xi(t) - t \cdot \Xi(1).$$

Все траектории  $\zeta(t)$  имеют первую производную, но не имеют второй,  $\zeta(0) = \zeta(1) = 0$ . Вычисляя коэффициенты разложения в ряд по той же системе, получаем, что плотность вероятности получающегося нормального распределения задается квадратичной формой  $-\frac{1}{2} \sum_{k=1}^{\infty} (k\pi)^4 a_k^2$ . Замена переменных, при которой квадратичная форма меры превращается в  $\frac{1}{2} \sum_{k=1}^{\infty} z_k^2$ , переведет квадратичную форму длины в  $\sum_{k=1}^{\infty} z_k^2 / (k\pi)^4$ .

Проводя аналогичные рассуждения в общем случае, получаем квадратичную форму длины в виде  $\sum_{k=1}^{\infty} z_k^2 / (k\pi)^{2\alpha}$ , где  $\alpha - 1/2$  — “гладкость” траекторий. Этот же результат можно сформулировать в терминах гладкости допустимого пространства линейных функционалов, поскольку каждый гладкий линейный функционал от функции будет линейным функционалом от ее первообразной.

В единичном гиперкубе все переменные равноправны относительно меры. Поэтому естественно принять, что масштаб по ним монотонно возрастает с увеличением номера  $n$  координаты, например как  $n^s$ . Будем строить метрики следующим образом. Сначала рассмотрим пра-норму  $\|\mathbf{y}\|$ , в которой все переменные равноправны, а потом изменим масштабы по формуле

$$\|\{x^{(1)}, \dots, x^{(n)}, \dots\}\| = \left\| \left\{ \frac{x^{(1)}}{\varphi(1)}, \dots, \frac{x^{(n)}}{\varphi(n)}, \dots \right\} \right\|.$$

Предположим, что пра-норма  $\|\{y^{(1)}, \dots, y^{(n)}, \dots\}\|$  конечна на некотором подпространстве пространства последовательностей  $\{y^{(1)}, \dots, y^{(n)}, \dots\}$  и обладает там следующими свойствами.

C<sub>1</sub>. Длина единичного отрезка на каждой оси координат равна 1:

$$\|\{0, \dots, 0, 1, 0, \dots\}\| = 1.$$

C<sub>2</sub>. Оси координат являются осями симметрии единичной сферы:

$$\|\{y^{(1)}, y^{(2)}, \dots, y^{(n)}, \dots\}\| = \|\{\pm y^{(1)}, \pm y^{(2)}, \dots, \pm y^{(n)}, \dots\}\|.$$

C<sub>3</sub>. Для любой перестановки натурального ряда

$$\begin{pmatrix} 1 & 2 & \dots & n & \dots \\ k_1 & k_2 & \dots & k_n & \dots \end{pmatrix}$$

биссектрисы координатных углов являются осями единичной сферы:

$$\|\{y^{(1)}, y^{(2)}, \dots, y^{(n)}, \dots\}\| = \|\{y^{(i_1)}, y^{(i_2)}, \dots, y^{(i_n)}, \dots\}\|.$$

C<sub>4</sub>.  $\|\{y^{(1)}, \dots, y^{(n)}, y^{(n+1)}, \dots\}\| = \lim_{n \rightarrow \infty} \|\{y^{(1)}, \dots, y^{(n)}, 0, 0, \dots\}\|$ .

Примером норм, обладающих свойствами C<sub>1</sub>–C<sub>4</sub>, могут служить нормы  $l_p$ . Однако они обладают еще свойством рекуррентности

$$P. \left\| \{y^{(1)}, y^{(2)}, y^{(3)}, \dots, y^{(n)}, \dots\} \right\| = \\ = \left\| \left\{ \left\| \{y^{(1)}, y^{(2)}, 0, 0, \dots\} \right\|, y^{(3)}, \dots, y^{(n)}, \dots \right\} \right\|,$$

т.е. единичная сфера в пространстве  $l_p$  определяется своим двумерным сечением.

Всюду далее мы предполагаем выполненными условия  $C_1, C_2, C_3, C_4$  (но не условие P), не оговаривая этого особо. Нам потребуются следующие два свойства пра-нормы.

ЛЕММА 1. Норма  $\|\mathbf{y}\|$  возрастает при увеличении модуля любой координаты  $y^{(k)}$ .

ЛЕММА 2. Сфера  $\|\mathbf{y}\| \leq r$  вписана в призматический гипероктаэдр  $\sum_{k=1}^n |y^{(k)}| \leq nd$ , касающийся сферы в точках  $\{\underbrace{\pm d, \pm d, \dots, \pm d}_n, 0, 0, \dots\}$ , где  $d \|\underbrace{\{1, \dots, 1, 0, \dots\}}_n\| = r$ .

Обе леммы доказываются элементарными геометрическими рассуждениями.

В качестве масштабов по координатам мы будем допускать функции  $\varphi(n)$  со следующими свойствами:

$$\varphi(n+1) \geq \varphi(n), \quad (M_1)$$

$$\varphi(kn) \geq \varphi(k) \varphi(n). \quad (M_2)$$

Оказывается, что в интересных случаях очень часто имеет место

УТВЕРЖДЕНИЕ. Существует такая константа  $C$ , зависящая от пра-нормы и функции  $\varphi$ , что

$$C \left\| \left\{ \underbrace{1, 1, \dots, 1}_n, 0, \dots \right\} \right\| \geq \varphi(n) \left\| \left\{ \underbrace{0, \dots, 0}_n, \frac{1}{\varphi(n+1)}, \dots, \frac{1}{\varphi(n+k)}, \dots \right\} \right\|. \quad (Y)$$

Так, для рекуррентных норм оно выполнено, если только диагональ единичного куба  $H$  в метрике  $\|\mathbf{x}\|$  является ограниченной. За константу  $C$  тогда можно принять  $\|\{1, 1, \dots, 1, 1, \dots\}\|$ . Оно выполнено также, если  $\sum_{k=1}^{\infty} 1/\varphi(k) < \infty$ ; тогда можно принять  $C = \sum_{k=1}^{\infty} 1/\varphi(k)$ .

Рассмотрим на гиперкубе  $H$  функции  $f(\mathbf{x})$ , удовлетворяющие условию Липшица по норме

$$\|\mathbf{x}\| = \left\| \left\{ \frac{x^{(1)}}{\varphi(1)}, \dots, \frac{x^{(n)}}{\varphi(n)}, \dots \right\} \right\|.$$

Пусть  $\mathcal{H}_1$  — класс всех функций на  $H$ , удовлетворяющих условию Липшица с коэффициентом 1. Выясним, каким может быть наилучший порядок остаточного члена в квадратурных формулах для функций класса  $\mathcal{H}_1$ . Нетрудно заметить, что таким же будет порядок остаточного члена для класса дифференцируемых функций, у которых норма первой производной ограничена единицей. Введем функцию

$$\Psi(n) = \prod_{k=1}^n \left[ \frac{\varphi(n)}{\varphi(k)} \right] = \left[ \frac{\varphi(n)}{\varphi(n-1)} \right]^n \cdot \Psi(n-1)$$

и “обратную” функцию  $\nu(N)$ ,  $\Psi(\nu(N)) \leq N$ ,  $\Psi(1 + \nu(N)) > N$ .

ТЕОРЕМА 2. Для класса функций  $\mathcal{H}_1$  при любом  $N$  существует квадратурная формула вида  $\sum_{i=1}^N \alpha_i f(\mathbf{x}_i)$  (где  $\sum_{i=1}^N \alpha_i = 1$ ,  $\alpha_i = 1/M$  при  $i = 1, \dots, M$ ,  $\alpha_i = 0$  при  $i > M$ ), дающая ошибку  $R_N$ :

$$|R_N| \leq \frac{\| \{ \overbrace{1, \dots, 1}^n, 0, \dots \} \|}{\varphi(n)} + \frac{1}{2} \left\| \left\{ \underbrace{0, \dots, 0}_n, \frac{1}{\varphi(n+1)}, \dots \right\} \right\|, \quad (4)$$

где  $n = \nu(N)$ .

СЛЕДСТВИЕ 2.1. Для функций класса  $\mathcal{H}_1$  (в смысле нормы  $\|\mathbf{x}\|$ ), обладающей свойством (У), существуют квадратурные формулы с остаточным членом порядка

$$\frac{1}{\varphi(n)} \cdot \| \{ \overbrace{1, \dots, 1}^n, 0, \dots \} \|,$$

где  $n = \nu(N)$ .

СЛЕДСТВИЕ 2.2. Для функций класса  $\mathcal{H}_1$  в смысле нормы  $l_p^s$ ,

$$\|\mathbf{x}\| = \left[ \sum_{k=1}^{\infty} \left( \frac{x^{(k)}}{k^s} \right)^p \right]^{1/p},$$

существует квадратурная формула с остаточным членом

$$|R_N| < n^{-s+1/p} \cdot \left[ 1 + 2^{-p} \frac{1}{sp-1} \right]^{1/p},$$

где  $n = \Pi[(1/s) \ln N]$ ,  $\Pi[\ ]$  — целая часть числа.

Доказательство теоремы 2. Пусть  $n = \nu(N)$ . Разобьем отрезок  $[0, 1]$  по каждой из первых  $n$  координатных осей на  $\Pi[\varphi(n)/\varphi(k)]$  равных отрезков, где  $k$  — номер координаты. Проведем через каждую точку деления гиперплоскость, перпендикулярную соответствующей оси. Они разобьют гиперкуб  $H$  на  $M = \prod_{k=1}^n \Pi[\varphi(n)/\varphi(k)]$  параллелепипедов. Очевидно,

$$M \leq N, \quad \Pi \left[ \frac{\varphi(n)}{\varphi(k)} \right] \geq \frac{1}{2} \frac{\varphi(n)}{\varphi(k)}.$$

Расстояние  $r$  по норме  $\|\mathbf{x}\|$  от центра параллелепипеда до наиболее удаленной точки — его вершины — есть

$$r = \left\| \left\{ \frac{u_1}{2}, \dots, \frac{u_n}{n}, \frac{1}{2}, \frac{1}{2}, \dots \right\} \right\|, \quad \text{где} \quad \frac{1}{u_k} = \Pi \left[ \frac{\varphi(n)}{\varphi(k)} \right].$$

По лемме 1

$$\begin{aligned} r &\leq \left\| \left\{ \frac{\varphi(1)}{\varphi(n)}, \dots, \frac{\varphi(n)}{\varphi(n)}, \frac{1}{2}, \frac{1}{2}, \dots \right\} \right\| \leq \\ &\leq \| \{ \overbrace{1, \dots, 1}^n, 0, 0, \dots \} \| + \frac{1}{2} \left\| \left\{ \underbrace{0, \dots, 0}_n, \frac{1}{\varphi(n+1)}, \dots \right\} \right\|. \end{aligned}$$

Замена  $f(\mathbf{x})$  в каждом параллелепипеде на значение  $f(\mathbf{x})$  в его центре даст в интеграле ошибку, не превосходящую  $r$  (так как константа Липшица — единица). Так как квадратурная формула по  $M$  точкам, взятым в центрах параллелепипедов с равными весами, дает точное значение для интеграла от ступенчатой функции, то  $|R_N| \leq r$ .

ТЕОРЕМА 3. *Какова бы ни была квадратурная формула по  $N$  точкам, точная на  $f(\mathbf{x}) = 0$ , среди функций класса  $\mathcal{H}_1$  для любого  $n$  найдется такая функция  $f$ , для которой ошибка  $R_N$*

$$|R_N| \geq \frac{1}{e} \frac{\|\{1, \dots, 1, 0, 0, \dots\}\|}{\varphi(n)} \left(1 - \frac{1}{n+1}\right) \left[\frac{1}{N} \Psi(n)\right]^{1/n}. \quad (5)$$

СЛЕДСТВИЕ 3.1. *Для функций класса  $\mathcal{H}_1$  (в смысле нормы  $\|\mathbf{x}\|$ ) не существует квадратурной формулы с остаточным членом, имеющим при всех  $N$  лучший порядок, чем*

$$\frac{1}{\varphi(n)} \|\{1, \dots, 1, 0, 0, \dots\}\|,$$

*и не существует формулы с остаточным членом, имеющим хотя бы при некоторых  $N$  лучший порядок, чем*

$$\frac{1}{\varphi(n+1)} \|\{1, \dots, 1, 0, 0, \dots\}\|,$$

где  $n = \nu(N)$ .

СЛЕДСТВИЕ 3.2. *Для функций класса  $\mathcal{H}_1$  в смысле нормы  $l_p^s$ ,*

$$\|\mathbf{x}\| = \left[ \sum_{k=1}^{\infty} \left( \frac{x^{(k)}}{k^s} \right)^p \right]^{1/p},$$

*не существует квадратурной формулы с остаточным членом  $R_N \lesssim 1/e n^{-S+1/p}$ ,  $n = \Pi[(1/3) \ln N]$ .*

Доказательство теоремы 3. Обозначим  $\|\mathbf{x}\|_n = \|\{x^{(1)}, \dots, x^{(n)}, 0, 0, \dots\}\|$ . Сконструируем плохую функцию в виде  $f(\mathbf{x}) = \min_{1 \leq i \leq n} \|\mathbf{x} - \mathbf{x}_i\|$ , где  $\mathbf{x}_i$  — узлы квадратурной формулы. Функция  $f(\mathbf{x})$  удовлетворяет условию Липшица с коэффициентом единица, так как в силу  $\|\mathbf{x}\|_n \leq \|\mathbf{x}\|$  и неравенства треугольника  $\|\mathbf{x}_n\|$  удовлетворяет этому же условию. Пусть  $V_n(r)$  — сфера  $\|\mathbf{x}\|_n \leq r$  (сферический цилиндр в  $H$ ) и  $|V_n(r)|$  — ее объем. Выберем  $|V_n(r)| = 1/N$ .

Сравнивая размеры линий уровня, получаем

$$\int_H f(\mathbf{x}) d\mu \geq N \int_{V_n(r)} \|\mathbf{x}\|_n d\mu = N \left(1 - \frac{1}{n+1}\right) r |V_n(r)| = r \left(1 - \frac{1}{n+1}\right).$$

Оценим теперь величину  $r$  снизу. Сфера  $V_n(r)$  является аффинным образом сферы  $\|\mathbf{y}\|_n \leq r$  объема  $|W_n(r)|$ .

$$|V_n(r)| = |W_n(r)| \cdot \prod_{k=1}^n \varphi(k).$$

Объем  $|W_n(r)|$  мы оценим сверху по лемме 2. Объем описанного призматического гипероктаэдра равен  $d^n n^n / n! \leq d^n e^n$ , где

$$r = \left\| \underbrace{\{d, \dots, d, 0, 0, \dots\}}_n \right\| = d \left\| \underbrace{\{1, \dots, 1, 0, 0, \dots\}}_n \right\|$$

Далее,

$$\frac{1}{N} \leq d^n e^n \cdot \prod_{k=1}^n \varphi(k) \cdot \prod_{k=1}^n \left[ \frac{\varphi(n)}{\varphi(k)} \right],$$

откуда, извлекая корень  $n$ -й степени, получим

$$r \geq \frac{1}{e} \frac{\left\| \underbrace{\{1, \dots, 1, 0, \dots\}}_n \right\|}{\varphi(n)} \cdot \left[ \frac{1}{N} \prod_{k=1}^n \frac{\varphi(n)}{\varphi(k)} \right]^{1/n}.$$

Если рассматривать  $\varphi(n)$  вида  $e^{-\alpha n}$ , то с точностью до множителей низшего порядка наш остаточный член на классе  $\mathcal{H}_1$  будет иметь порядок  $e^{-c\sqrt{\ln n}}$ . Вообще, какой бы вид ни имела функция  $\varphi(n)$ , остаточный член в классе  $\mathcal{H}_1$  будет убывать медленнее любой степени  $N^{-\beta}$ . Это следует из теоремы о том, что остаточный член для  $\alpha$  раз дифференцируемых функций на  $s$ -мерном кубе убывает не быстрее  $N^{-\alpha/s}$ . Чтобы получить степенной порядок убывания, надо существенно сузить рассматриваемый класс функций и потребовать существования у функций класса каких-либо старших производных (причем сколь угодно высокого порядка).

## О МОДИФИЦИРОВАННОМ РЕКУРРЕНТНОМ АЛГОРИТМЕ КОРОБОВА ПОСТРОЕНИЯ РАВНОМЕРНОЙ СЕТКИ В МНОГОМЕРНОМ КУБЕ <sup>\*)</sup>

При решении многих кубатурных задач требуется строить равномерные последовательности точек в единичном многомерном кубе, проекции которых на грани меньших размерностей также являются достаточно равномерными последовательностями. В известной монографии Н.М. Коробова [1] были предложены решетки такого типа, позволяющие строить оптимальные квадратурные формулы для некоторых классов гладких функций.

Однако для построения этих решеток требуется предварительно вычислить оптимальные коэффициенты, задающие основной репер решетки. Предварительных вычислений требуют и изученные С.Л. Соболевым решетки, связанные с

<sup>\*)</sup> Вопросы вычислительной и прикладной математики. Вып. 38. — Ташкент, 1970. — С. 127–133 (совм. с Е.Б. Вул).

плотнейшими упаковками шаров [2, лекция 20]. Поэтому иногда практически удобнее не гнаться за оптимальностью решетки, а воспользоваться простыми рекуррентными формулами.

Излагаемый ниже алгоритм хорошо зарекомендовал себя в многочисленных расчетах [3, 4]. Он является модификацией алгоритма, предложенного Н.М. Коробовым [5] в 1957 г.

Решетка задается двумя параметрами — целыми числами  $P$  и  $g$ . Число  $P$  должно быть лежандровым простым числом, т.е. простым числом вида  $P = 2Q + 1$ , где  $Q$  — также простое число. Число  $g$  — так называемый первообразный корень по модулю  $P$ . Для лежандрова простого числа им является любой квадратичный невычет, например, им всегда будут числа  $P - 3^m$  и  $P - 4^m$ .

Приведем небольшую таблицу лежандровых простых чисел.

Таблица лежандровых чисел

107	6599	20183	220919	2000303
587	8039	59983	400067	10000223
2027	9839	40127	400307	10000643
2039	10007	40499	600983	10000763
2063	10079	50147	864119	10000799
5087	10103	100043	875363	
5099	20123	100103	999959	

За  $n$ -ю координату  $\xi_n^{(k)}$   $k$ -й точки  $S$ -мерного куба принимается

$$\xi_n^{(k)} = \text{Др} \left\{ \frac{(g^k)^n}{P} \right\} = \text{Др} \left\{ \frac{g^{kn}}{P} \right\}, \quad (1)$$

где  $\text{Др} \{x\}$  — дробная часть числа  $x$ .

Как известно из курса теории чисел, эти величины легко сосчитать рекуррентно по правилам

$$\text{Др} \left\{ \frac{a}{p} \right\} = \frac{\text{Ост}_P(a)}{P}, \quad (2)$$

$$\text{Ост}_P\{ab\} = \text{Ост}_P[\text{Ост}_P\{a\} \cdot \text{Ост}_P\{b\}], \quad (3)$$

так как переход  $\text{Ост}_P$  от натурального числа  $a$  к остатку от деления  $a$  на  $P$  (в целых числах) есть гомоморфизм кольца натуральных чисел в поле остатков по модулю  $P$ .

Поэтому при рекуррентном построении  $k$ -го узла решетки, кроме чисел  $P$  и  $g$ , в памяти держится еще число  $g_k = \text{Ост}_P\{g^k\}$ , и рекуррентно вычисляются  $g_{kn} = \text{Ост}_P\{(g^k)^n\}$ , а сами координаты досчитываются как  $\xi_n^{(k)} = g_{kn}/P$ . При переходе к построению  $k + 1$ -го узла надо досчитать

$$g_{k+1} = \text{Ост}_P\{g_k \cdot g\}.$$

Для удобства наш алгоритм реализован в виде процедуры на языке АЛГОЛ. Обращение к процедуре для получения первого случайного числа 1-й серии имеет вид `rand (1)`, для получения чисел внутри любой серии — `rand (3)`, для получения первого числа последующих серий — `rand (2)`. Заметим, что  $p, g, g_n, g_k, sl$  — глобальные переменные.

Приводим описание процедуры:

```

procedure rand(m); value m;
integer m; begin switch s:=m1,m2,m3;
go to s[m];
m1:gk:=1;
m2:gn:=1;
m3:gn:=gn*gk-entier((g*gk)/p)*p;
s1:=gn/p end;

```

Числа  $\xi_n^{(2k+1)}$  ( $n = 1, 2, \dots$ ) при  $0 < 2k + 1 < Q$  и  $Q < 2k + 1 < P$  состоят из всех дробей вида  $m/P$ ,  $m = 1, \dots, P - 1$ . При больших  $P$  и  $g$  порядок следования этих чисел весьма нерегулярен. Числа  $\xi_n^{(2k)}$ ,  $n = 1, 2, \dots$ , при  $0 < 2k < P - 1$  пробегает в нерегулярном порядке все дроби вида  $m_1/P$ , где  $m_2 = \text{Oстр}_P(m^2)$ ,  $m = 1, \dots, P - 1$ . Поэтому такие последовательности можно считать псевдослучайными.

Период каждой нечетной серии равен  $P - 1 = 2Q$  и для четной серии  $-Q$ . Серия  $\xi_n^{(0)}$  является последовательностью  $(P - 1)/P, 1/P, (P - 1)/P$  периода, равного двум. Серия  $\xi_n^{(P-1)}$  состоит из одних  $1/P$ . Далее серии повторяются с периодом  $P - 1$ . Заметим, что по построению  $\xi_n^{(k)} = \xi_k^n$ ; эту симметрию следует иметь в виду. В частности, ею можно воспользоваться для проверки работы алгоритма на машине.

В работе [5] доказано, что при небольших размерностях точки  $\vec{\xi}_1, \dots, \vec{\xi}_{P-1}$  распределены приблизительно равномерно с флуктуациями  $O(1/\sqrt{P})$ . Представляется правдоподобным, что это справедливо для более коротких последовательностей

$$\vec{\xi}_1, \dots, \vec{\xi}_N, \quad N < P - 1.$$

Стоит отметить, что алгоритм Коровова позволяет строить кубатурные формулы и для интегрирования по счетномерному тору [7], достаточно точные для некоторых специальных классов функций; ср. [7, 8].

**ТЕОРЕМА.** Пусть на единичном счетномерном торе  $H$ , точками которого являются последовательности  $x : x_1, x_2, \dots, x_i, \dots$ ,  $0 < x_i < 1$ ,  $i = 1, 2, \dots$ , задана функция  $f(x_1, x_2, \dots)$ . Если существует такая константа  $C > 0$ , что для каждого набора  $s_1 < s_2 < s_3 < \dots < s_\nu = s$

$$\sup_{x \in H} \left| \frac{\partial^{2\nu} f}{\partial x_{s_1}^2 \dots \partial x_{s_\nu}^2} \right| \leq C^s \frac{1}{s!}, \quad (4)$$

то при  $\nu = P - 1$

$$\left| \int_H f(x_1, x_2, \dots) \prod_{s=1} dx_s - \frac{1}{N} \sum_{k=1}^N f(\xi_1^{(k)}, \xi_2^{(k)}, \dots) \right| = O\left(\frac{1}{\sqrt{N}}\right), \quad (5)$$

где координаты  $\xi_n^{(k)}$  заданы формулой (1).

**Доказательство.** Так как функция периодична, то из оценки (4) следует оценка для коэффициентов Фурье при  $m_j \neq 0$ :

$$|C(m_1, \dots, m_\nu)| \leq C^s \frac{1}{\nu!} \prod_{s=1}^\nu (|m_s| + \delta_s)^{-2}, \quad (6)$$

$C(m_1, \dots, m_\nu, 0, 0, \dots) \delta_s = 1$  при  $m_s = 0$  и  $\delta_s = 0$  при  $m_s \neq 0$ .

Обозначим  $f_0 = C(0, 0, \dots)$ ,

$$\begin{aligned} f_\nu(x_1, \dots, x_\nu, \dots) &= f_\nu(x_1, \dots, x_\nu, 0, 0, \dots) = \\ &= \sum_{\substack{m_s = -\infty \\ m_\nu \neq 0}}^{\infty} C(m_1, \dots, m_\nu) e^{2\pi i(m_1 x_1 + \dots + m_\nu x_\nu)}. \end{aligned}$$

Легко видеть, что

$$f(x_1, \dots, x_n, \dots) = \sum_{\nu=0}^{\infty} f_\nu(x_1, \dots, x_\nu).$$

Обозначим через  $I[\cdot]$  оператор интегрирования по гиперкубу  $H$  по мере  $dV = \prod dx_s$ , через  $S[f]$  — квадратурную сумму  $\frac{1}{N} \sum_{k=1}^N f(x^{(k)})$  и через  $R[\cdot] = I[\cdot] - S[\cdot]$ . Тогда

$$\begin{aligned} |I[f] - S[f]| &= |R[f]| \leq \sum_{\nu=1}^{\infty} |R[f_\nu]| = \sum_{\nu=1}^N |R[f_\nu]| + \sum_{\nu=N+1}^{\infty} |R[f_\nu]| \leq \\ &\leq \sum_{\nu=1}^N \left\{ \frac{\nu-1}{\sqrt{N}} \sum_m |c(m_1, \dots, m_\nu)| + \frac{7\nu C}{N} \right\} + \sum_{\nu=N+1}^{\infty} 2 \sup_x |f_\nu(x)|, \end{aligned}$$

где первое слагаемое оценено по теореме 2 работы [5]. С помощью (6) получаем оценки

$$\sup_x |f_\nu(x)| = O\left(\frac{1}{\nu!}\right); \quad \sum_{N+1}^{\infty} 2 \sup_x |f_\nu(x)| = O(N^{-2}).$$

Оценивая из тех же соображений первое слагаемое, выводим

$$\sum_{\nu=1}^N \left\{ \frac{\nu-1}{\sqrt{N}} \sum_m |c(m_1, \dots, m_\nu)| + \frac{7\nu C}{N} \right\} \leq \sum_{\nu=1}^N \left\{ c^\nu \frac{\nu Q^\nu}{\nu!} \frac{1}{\sqrt{N}} + \frac{7\nu C}{\nu!} \frac{1}{N} \right\} = O\left(\frac{1}{\sqrt{N}}\right),$$

что и требовалось доказать.

Отметим, что класс функций, удовлетворяющих условию (4), не является чрезмерно усложненным. Например, ему удовлетворяет функция, заданная рядом

$$f(x_1, \dots, x_n, \dots) = \sum_{s=0}^{\infty} \frac{t^s}{s!} K^s g(y),$$

где:

$K^s$  —  $s$ -я итерация интегрального оператора с ядром  $K(y, x)$  на единичном квадрате, периодическим по каждому переменному и имеющим ограниченные производные вплоть до смешанной производной второго порядка по каждому переменному;

функция  $g(y)$  также предполагается периодической и дважды непрерывно дифференцируемой;

$t$  — любое действительное число.

Следует отметить, что многие предлагавшиеся алгоритмы построения равномерно распределенных последовательностей (или псевдослучайных чисел) таковыми не являются. Так, в книге Голенко [6] предложена последовательность псевдослучайных чисел  $\eta_n$ ,  $0 \leq \eta_n \leq 1$ , почти равномерно распределенных на отрезке  $[0,1]$ . Но, как отмечает автор работы [6], пары  $(\xi_n^{(1)}, \xi_n^{(2)})$ , где  $\xi_n^{(1)} = \eta_{2n-1}$ ,  $\xi_n^{(2)} = \eta_{2n}$ , распределены в квадрате уже неравномерно. Наш алгоритм свободен от этого недостатка по построению. Правда, он несколько длинен, поэтому им разумно пользоваться только тогда, когда от расчетов требуется существенная точность.

Для демонстрации точности алгоритма опишем результаты следующего модельного расчета, связанного с вычислением интеграла от кусочно постоянных функций. Поскольку это безразлично, проще здесь говорить не об интеграле от некоторой сложно задаваемой функции на  $n$ -мерном кубе, а о среднем значении некоторого функционала от траектории, которая разыгрывается по  $n$  псевдослучайным числам.

Разыгрывались случайные пути на кубической решетке. Длина пути равнялась 250, общее число полученных путей — 720. Выбор следующего шага делался с вероятностью  $1/6$  для каждого возможного продолжения. У полученных ломаных исследовалось среднее арифметическое квадрата расстояния  $h^2$  между начальной и конечной точками и гистограмма распределения этой величины. Теоретические значения для математического ожидания и дисперсии:  $\mathbf{M}h^2 = n = 250$  и  $\mathbf{D}h^2 = 2n(n-1)/3 = (2/3) \cdot 249 \cdot 250$ .

Таким образом, ожидаемое среднее должно быть близко к 250 с разбросом, определяемым величиной  $\alpha = \sqrt{\mathbf{D}h^2/750} = 7,59$ . Счет дал среднее значение  $\langle h^2 \rangle$ , равное 244,31.

Для более полной обработки результатов был использован критерий согласия  $\chi^2$ . Возможные значения квадрата расстояния  $h^2$  были разбиты на 6 групп:

- |                               |                               |                               |
|-------------------------------|-------------------------------|-------------------------------|
| 1. $h^2 < 112,5$ ;            | 2. $112,5 \leq h^2 < 237,5$ ; | 3. $237,5 \leq h^2 < 362,5$ ; |
| 4. $362,5 \leq h^2 < 487,5$ ; | 5. $487,5 \leq h^2 < 612,5$ ; | 6. $612,5 \leq h^2$ .         |

Пусть  $\nu_i$ , где  $i = 1, \dots, 6$ , — число ломаных, у которых квадрат расстояния попал в  $i$ -ю группу. Критерий  $\chi^2$  имеет вид

$$\chi^2 = \sum_{i=1}^6 \frac{\nu_i^2}{N P_i} - N,$$

где  $P_i$  — теоретическая вероятность того, что величина  $h^2$  попала в  $i$ -ю группу. Найденное значение  $\chi^2$  оказалось равным 4,48, и оно близко к 50%-ному уровню, что вполне удовлетворительно.

## СПИСОК ЛИТЕРАТУРЫ

1. Коробов Н.М. Теоретикочисловые методы в приближенном анализе. — М.: Физматгиз, 1963.
2. Соболев С.Л. Лекции по теории кубатурных формул. — Новосибирск: НГУ, 1965.
3. Гельфанд И.М. и др. // Изв. вузов. Сер. матем. 1958. № 5 (6).
4. Введенская Н.Д. // Проблемы передачи информации. 1968. Вып. 3.

5. Коробов Н.М. // Докл. АН СССР. 1957. Т. 115, № 6.
6. Голенко Д.И. Моделирование и статистический анализ псевдослучайных чисел на электронных вычислительных машинах. — М.: Наука, 1965.
7. Ченцов Н.Н. // ЖВМиМФ. 1961. Т. 1, № 3.
8. Соболев И.М. // ЖВМиМФ. 1961. Т. 1, № 5.

## ОБ ОЦЕНИВАНИИ ЗАДАННОЙ В СЛУЧАЙНЫХ УЗЛАХ ФУНКЦИИ МЕТОДОМ НАИМЕНЬШИХ КВАДРАТОВ<sup>\*)</sup>

При вычислениях по методу Монте-Карло часто удается получить значения гладкой функции  $f(x)$  для некоторого набора случайных аргументов. Возникает задача оценивания этой функции (или же ее моментов — интегралов  $\int fh dx$ ) по “наблюденным” значениям.

Пусть  $X^n \subset \mathbb{R}^n$  —  $n$ -мерный параллелепипед  $a_i \leq x_i \leq b_i$ ,  $i = 1, \dots, n$ , объем которого, не ограничивая общности, полагаем равным единице,  $T = (t_1, \dots, t_n)$ ,  $W_p^T(X^n)$  — класс функций, у которых по каждой переменной  $x_i$  существует обобщенная в норме  $L_p(X^n)$  производная порядка  $t_i$ . Ошибка лучшего на классе  $W_p^T(X^n)$  способа оценивания функции по  $N$  ее значениям убывает как  $N^{-\rho}$ ,  $\rho = \left(\sum_{i=1}^n t_i^{-1}\right)^{-1}$  при  $p\rho > 1$  [1]. Для рандомизированных способов интегрирования функций из  $W_p^T(X^n)$  оценка снизу порядка вероятной ошибки равна  $N^{-\rho-1/2}$ , в то время как для неслучайных способов интегрирования такая оценка равна  $N^{-\rho}$  [2]. Проблема состоит в построении оценок функции  $f \in W_p^T(X^n)$  и ее моментов с оптимальными порядками вероятных погрешностей при весьма общих предположениях о вероятностном законе распределения узлов наблюдений.

Если  $f \in W_p^T(X^n)$ , то лучший порядок убывания  $L_p(X^n)$ -остатка можно получить при разложении  $f(x)$  по системе кусочно многочленных функций  $\varphi_{ki}(x)$ ,  $k = 1, \dots, q$ ,  $i = 1, \dots, m$ , заданных на непересекающихся подынтервалах  $X_k^n$ ,  $\bigcup_k X_k^n = X^n$ , так, что  $\varphi_{ki}(x) = 0$  при  $x \notin X_k^n$  и  $\int \varphi_{ki}\varphi_{lj} dx = \delta_{ki}\delta_{ij}$ . Если функции  $\varphi_{ki}$  представляют собой ортонормированные на  $X_k^n$   $n$ -мерные многочлены Лежандра, то можно подобрать такие разбиения интервала  $X^n \int_{X_k^n} dx \asymp q$  и такое конечное  $m$ , зависящее от  $T$ , что с ростом  $M = mq$

$$\sup_{f \in W_p^T(X^n)} \left\| f(x) - \sum_{k=1}^q \sum_{i=1}^m \left( \int f \varphi_{ki} dx \right) \varphi_{ki}(x) \right\|_p = O(M^{-\rho}) \quad (1)$$

при  $p \geq 2$ ,  $p\rho > 1$  [1].

<sup>\*)</sup> Теория вероятн. и ее примен. 1990. Т. 35, вып. 4. С. 771–774 (совм. с А.И. Корякиным).

Пусть заданы  $N$  значений функции  $f(x)$  в случайных узлах  $x^j \in X^n$ ,  $j = 1, \dots, N$ , независимо распределенных с известной плотностью вероятности  $g(x)$ , причем  $g^{-1} \in L_\infty(X^n)$ . Для вычисления по этим значениям вектора коэффициентов разложения  $C = \int f \Phi dx$ ,  $\Phi = (\varphi_{11}, \dots, \varphi_{qm})^T$ , функции  $f(x)$  будем использовать оценку  $\hat{C}(M)$  [3]. Такая оценка представляет собой решение по методу наименьших взвешенных квадратов системы  $N$  линейных уравнений

$$\sum_{k=1}^q \sum_{i=1}^m \hat{C}_{ki} \varphi_{ki}(x^j) = f(x^j)$$

относительно  $M$  неизвестных  $\hat{C}_{ki}$  и имеет вид

$$\hat{C}(M) = \lim_{\delta \rightarrow 0} (A^T B A + \delta^2 I)^{-1} A^T B F.$$

Здесь  $A = (\Phi(x^1), \dots, \Phi(x^N))^T$ ,  $F = (f(x^1), \dots, f(x^N))^T$ ,  $B$  — весовая диагональная матрица с элементами  $g^{-1}(x^j)$ ,  $j = 1, \dots, N$ , выбранная так, чтобы матрица

$$N^{-1} A^T B A = N^{-1} \sum_{j=1}^N \Phi(x^j) \Phi^T(x^j) g^{-1}(x^j) \quad (2)$$

сходилась по вероятности при  $N \rightarrow \infty$  к единичной матрице  $I$ .

Проекционную оценку функции  $f(x)$  получаем по ее значениям в случайных узлах в виде  $\Phi^T(x) \hat{C}(M)$ . Если для  $h(x)$  известны интегралы  $\int h \varphi_{ki} dx$ , то, выделив главную часть функции  $f(x) \approx \Phi^T(x) \hat{C}(M)$ , получаем оценку интеграла  $\int_{X^n} f h dx$  в виде

$$\hat{C}_h(M) = \hat{C}^T(M) \int \Phi h dx + N^{-1} \sum_{j=1}^N h(y^j) [f(y^j) - \Phi^T(y^j) \hat{C}(M)] v^{-1}(y^j). \quad (3)$$

Здесь  $y^j$ ,  $j = 1, \dots, N$ , — новые случайные векторы, независимо распределенные в  $X^n$  с плотностью  $v(y)$ , причем  $h^2(y) v^{-1}(y) \in L_\infty(X^n)$ .

Ошибка оценки  $\hat{C}(M)$  коэффициентов разложения  $C$  определяется убывающим с ростом  $M$  остатком разложения  $\|f - C^T \Phi\|_2$  [3]

$$\|\hat{C}(M) - C\| = \|(A^T B A)^{-1} A^T B F - C\| \leq \|N(A^T B A)^{-1}\| \cdot \|N^{-1} A^T B (F - AC)\|, \quad (4)$$

если норма матрицы  $N(A^T B A)^{-1}$  ограничена не слишком большой величиной. Для этого в каждый  $X_k^n$  должно попасть не менее  $t$  “хорошо” расположенных точек  $x^j$ , что ограничивает число  $q$  подынтервалов  $X_k^n$  и, следовательно, точность оценок  $\Phi^T(x) \hat{C}(M)$ ,  $\hat{C}_h(M)$ . Оценку допустимого значения  $q$  при заданном числе  $N$  случайных узлов, распределенных независимо от способа разбиения  $X^n$ , дает следующая теорема.

**ТЕОРЕМА 1.** Пусть плотность  $g(x)$  распределения независимых случайных узлов ограничена снизу:

$$\text{vrai inf}_{x \in X^n} g(x) = K_1^{-1}, \quad 1 \leq K_1 < \infty; \quad (5)$$

объемы  $V_k$  подынтервалов разбиения  $X_k^n$  не слишком малы:

$$V_k \geq (aq)^{-1}, \quad 1 \leq a < \infty, \quad k = 1, \dots, q; \quad (6)$$

а базисные ортонормированные функции относительно равномерно ограничены:

$$|\varphi_{ki}(x)| \leq K_2 V_k^{-1/2}, \quad 1 \leq K_2 < \infty, \quad i = 1, \dots, m. \quad (7)$$

Если  $q$  выбрано так, что

$$q < N \ln^{-1} N (1 - K^{-1})^2 (m^2 K_1^2 K_2^4 a \varepsilon^3)^{-1}, \quad (8)$$

то при  $N \geq \varepsilon^{-1}$ ,  $0 < \varepsilon < 1$  и любом  $K > 1$

$$\mathbf{P}\{\|N(A^T B A)^{-1}\| \leq K\} \geq 1 - \varepsilon. \quad (9)$$

Доказательство. Пусть  $\nu_i$ ,  $i = 1, \dots, M$ , — собственные значения неотрицательно определенной матрицы  $N^{-1} A^T B A$  блочно-диагонального вида (2),  $\nu_1 = \min \nu_i$ . Норма обратной матрицы

$$\|N(A^T B A)^{-1}\| = \nu_1^{-1}$$

ограничена, т.е.  $\nu_1^{-1} \leq K$ , если  $|1 - \nu_1| \leq 1 - K^{-1}$ , и, тем более, если

$$\sum_i (\nu_i - 1)^{2l} \leq (1 - K^{-1})^{2l},$$

$l$  целое. Поэтому из обобщенного неравенства Чебышева получаем

$$\mathbf{P}\{\|N(A^T B A)^{-1}\| \leq K\} \geq 1 - (1 - K^{-1})^{-2l} \mathbf{M} \sum_{i=1}^M (\nu_i - 1)^{2l}. \quad (10)$$

Для независимых случайных узлов  $x^j$

$$\begin{aligned} \mathbf{M} \sum_{i=1}^M (\nu_i - 1)^{2l} &= \mathbf{M} \operatorname{Sp} (N^{-1} A^T B A - I)^{2l} = \\ &= \mathbf{M} \sum_{k=1}^q \sum_{i_1=1}^m \dots \sum_{i_{2l}=1}^m \left[ N^{-1} \sum_{j_1}^N \varphi_{k i_1}(x^{j_1}) \varphi_{k i_2}(x^{j_1}) g^{-1}(x^{j_1}) - \delta_{i_1 i_2} \right] \dots \\ &\quad \dots \left[ N^{-1} \sum_{j_{2l}=1}^N \varphi_{k i_{2l}}(x^{j_{2l}}) \varphi_{k i_1}(x^{j_{2l}}) g^{-1}(x^{j_{2l}}) - \delta_{i_1 i_{2l}} \right] = \\ &= \sum_{k=1}^q \sum_{i_1, \dots, i_{2l}=1}^m N^{-2l} \sum_{j_1, \dots, j_{2l}=1}^N \int_{(X^n)^N} \dots \int [\varphi_{k i_1}(x^{j_1}) \varphi_{k i_2}(x^{j_1}) g^{-1}(x^{j_1}) - \delta_{i_1 i_2}] \times \dots \\ &\quad \dots \times [\varphi_{k i_{2l}}(x^{j_{2l}}) \varphi_{k i_1}(x^{j_{2l}}) g^{-1}(x^{j_{2l}}) - \delta_{i_1 i_{2l}}] g(x^1) \dots g(x^N) dx^1 \dots dx^N. \quad (11) \end{aligned}$$

Каждый из интегралов  $J$ , стоящих в последнем выражении, определяется индексами  $k; i_1, \dots, i_{2l}; j_1, \dots, j_{2l}$ , причем “слово”  $j_1, \dots, j_{2l}$  записано алфавитом из  $N$  “букв”  $n = 1, \dots, N$ . Пусть  $n_1, \dots, n_s$  — перечень всех “букв”, встречающихся в некотором “слове”  $j_1, \dots, j_{2l}$ . Тогда соответствующий этому “слову” интеграл  $J(j_1, \dots, j_{2l})$  равен произведению  $s$  интегралов по переменным  $x^{n_1}, \dots, x^{n_s}$  от зависящих от этих переменных частей подынтегрального произведения (11), так как интегрирование по  $x^j$ ,  $j \notin (n_1, \dots, n_s)$ , дает  $\int g(x^j) dx^j = 1$ . Если какая-либо “буква” встречается в “слове” ровно один раз, то соответствующие интегралы  $J$  равны нулю в силу ортонормированности функций  $\varphi_{ki}(x)$ . Таким образом, в (11) вклад дают только интегралы, соответствующие “словам”, в которых каждая встречающаяся “буква” встречается не менее двух раз. В этом случае  $s \leq l$ .

Пусть  $d(2l; n_1, \dots, n_s)$  — множество всех “слов” длины  $2l$ , записанных  $s$  “буквами”  $n_1, \dots, n_s$  из алфавита  $\{1, \dots, N\}$  так, что каждая “буква” встречается не менее двух раз. Тогда

$$\begin{aligned} \text{Card} \bigcup_{\{n_1, \dots, n_s\}} d(2l; n_1, \dots, n_s) &= C_N^s \text{Card} d(2l; 1, \dots, s) \leq \\ &\leq N(N-1) \dots (N-s+1) s^{2l} (s!)^{-1} \leq N^s s^{2l-s} e^s, \end{aligned} \quad (12)$$

поскольку все подалфавиты  $\{n_1, \dots, n_s\}$  алфавита  $\{1, \dots, N\}$  равноправны, а число подалфавитов равно числу сочетаний. Число же интересующих нас “слов” в подалфавите не превосходит числа всех “слов” длины  $2l$  в этом подалфавите. (Можно выписать оценку снизу для этого числа “слов”, которая показывает, что оценка сверху  $s^{2l}$  не может быть существенно улучшена.)

По условию теоремы (5), (7) при  $j_1, \dots, j_{2l} \in d(2l; n_1, \dots, n_s)$

$$J(k; i_1, \dots, i_{2l}; j_1, \dots, j_{2l}) \leq (K_1 K_2^2)^{2l} V_k^{s-2l}.$$

Отсюда и из (11), (12) получаем

$$\begin{aligned} \mathbf{M} \sum_{i=1}^M (\nu_i - 1)^{2l} &\leq N^{-2l} \sum_{k=1}^q \sum_{i_1, \dots, i_{2l}=1}^m \sum_{s=1}^l \sum_{s=1 \cup d(2l; n_1, \dots, n_s)} J(k; i_1, \dots, i_{2l}; j_1, \dots, j_{2l}) \leq \\ &\leq (K_1 K_2^2 m)^{2l} \sum_{k=1}^q \sum_{s=1}^l (s N^{-1} V_k^{-1})^{2l-s} e^s. \end{aligned}$$

Последнее выражение можно оценить без существенной потери точности, если принять  $N^{l-s} \geq (s V_k^{-1})^{l-s}$ ,  $s = 1, \dots, l$ . Тогда

$$\begin{aligned} \mathbf{M} \sum_{i=1}^M (\nu_i - 1)^{2l} &\leq (K_1 K_2^2 m)^{2l} \sum_{k=1}^q (l N^{-1} V_k^{-1})^l e^l \leq \\ &\leq (K_1^2 K_2^4 m^2 a e)^l N^{-l} (l q)^{l+1} = (K_3 l M)^{l+1} N^{-l}, \end{aligned}$$

$$K_3 = (K_1^2 K_2^4 a e)^{l/(l+1)} m^{(l-1)/(l+1)} > 1.$$

Отсюда и из (10) следует, что неравенство  $v_1^{-1} \leq K$  выполняется с вероятностью, не меньшей  $1 - \varepsilon$ , при

$$M \leq (\varepsilon N^{-1})^{1/(l+1)} N(K_3 l)^{-1} (1 - K^{-1})^{2l/(l+1)}. \quad (13)$$

Возьмем  $l = \lceil \ln N \rceil$ , что обеспечивает минимальное с точностью до не зависящей от  $N$  константы значение правой части (13), и  $N \geq \varepsilon^{-1}$ . Тогда

$$\varepsilon^{1/(l+1)} \geq N^{-1/(l+1)} \geq e^{-\ln N / (\ln N + 1)} \geq e^{-1},$$

и неравенство (13), обеспечивающее утверждение теоремы (9), выполняется при выполнении неравенства (8).

Доказанная теорема дает, кроме прочего, следующий полезный теоретико-вероятностный результат. Для того чтобы с вероятностью, сходящейся к единице, в каждый из подынтервалов  $X_k^n$ ,  $k = 1, \dots, q$ ,  $q \rightarrow \infty$ , объемы которых  $V_k \asymp q^{-1}$ , попало по крайней мере по  $m$  точек, достаточно распределять на  $\bigcup_k X_k^n$  такое число  $N$  независимых случайных точек с плотностью  $g(x)$ ,  $\text{vrai inf } g(x) > 0$ , чтобы  $q = o(m^2 N \ln^{-1} N)$ .

**ТЕОРЕМА 2.** Пусть число подынтервалов разбиения  $q$  растет как  $N \ln^{-1} N$  по порядку величины, удовлетворяя при этом неравенству (8).

1°. Для функции  $f(x) \in W_{2p}^T(X^n)$ ,  $p \in \mathbb{N} \cup \infty$ ,  $pr > 1/2$ , уклонение проекционной оценки  $\Phi^T(x) \hat{C}(M)$  от  $f(x)$  в норме  $L_{2p}(X^n)$  имеет по всей вероятности порядок  $N^{-\rho} \ln^\rho N$ .

2°. Если функция  $h(x) \in L_\infty(X^n)$  известна и  $f(x) \in W_2^T(X^n)$ ,  $\rho > 1/2$ , то ошибка построенной по двойной выборке оценки  $\hat{C}_h(M)$  (3) интеграла

$$\int_{X^n} fh dx$$

имеет по вероятности порядок  $N^{-\rho-1/2} \ln^\rho N$ .

Доказательство основано на оценивании математического ожидания случайной величины

$$G_{2p} = \sum_k \left( \sum_i |\hat{C}_{ki}(M) - C_{ki}|^2 \right)^p, \quad p \in \mathbb{N},$$

при условии выполнения события  $H = \{\|N(A^T B A)^{-1}\| \leq K\}$ , когда

$$G_{2p} \leq K^{2p} \sum_{k=1}^q \left( \sum_{i=1}^m \left| N^{-1} \sum_{j=1}^N \varphi_{ki}(x^j) [f(x^j) - C^T \Phi(x^j)] g^{-1}(x^j) \right|^2 \right)^p. \quad (14)$$

При оценивании условного математического ожидания используется неравенство

$$\mathbf{M}[\xi^2 | H] \leq P^{-1} \{H\} \mathbf{M} \xi^2,$$

справедливое для неотрицательной случайной величины  $\xi^2$ , и примененный при доказательстве теоремы 1 способ вычисления математических ожиданий произведений  $2p$  сумм по  $j = 1, \dots, N$ . При выполнении неравенств (5)–(7) оценка

$\mathbf{M}[G_{2p}|H]$  дает возможность получить, что при выполнении события  $H$  с вероятностью  $1 - \varepsilon > 1/2$  выполняется одно из соотношений

$$\|f - \Phi^T \widehat{C}(M)\|_{2p} \leq \|f - C^T \Phi\|_{2p} [1 + \mathbf{P}^{-1}\{H\} \varepsilon^{-1/2} O(pqN^{-1})^{1/2}], \quad p \in \mathbb{N},$$

$$\|f - \Phi^T \widehat{C}(M)\|_{\infty} \leq \|f - C^T \Phi\|_{\infty} + \mathbf{P}^{-1}\{H\} \varepsilon^{-1/2} \|f - C^T \Phi\|_{2 \ln N} O(qN \ln N^{-1})^{1/2}.$$

По теореме 1 при выполнении неравенства (8) вероятность одновременного выполнения  $H$  и одного из двух последних соотношений — не менее  $1 - 2\varepsilon$ . Следовательно, первое утверждение теоремы имеет место, если выбирать  $q \asymp N \ln^{-1} N$  так, чтобы выполнялось (8), поскольку функции  $\varphi_{ki}(x)$ , для которых справедливо соотношение (1), удовлетворяют неравенствам (6), (7).

Так же получаем второе утверждение теоремы, используя оценку  $\mathbf{M}(G_{2p}|H)$  (14), так как при  $h^2 v^{-1} \in L_{\infty}(X^n)$

$$\begin{aligned} \mathbf{M}_{x,y} \left( \left( \widehat{C}_h(M) - \int_{X^n} fh \, dx \right)^2 \middle| H \right) &= \\ &= \mathbf{M}_x \left( N^{-1} \int_{X^n} h^2 \{ \Phi^T (\widehat{C}(M) - C) + (f - C^T \Phi) \}^2 v^{-1} dy \middle| H \right) = \\ &= N^{-1} O(\mathbf{M}_x [\|\widehat{C}(M) - C\|^2 | H] + \|f - C^T \Phi\|_2^2). \end{aligned}$$

Полученные порядки убывания вероятных погрешностей оценивания функций и их моментов лишь на логарифмический множитель отличаются от оптимальных на классах  $W_p^T(X^n)$  порядков. Оптимальные порядки можно получать с помощью рассматриваемых оценок, если специальным образом распределять случайные узлы так, чтобы в каждый подынтервал  $X_k^n$  заведомо попадало по  $N_k > m^2$  узлов,  $k = 1, \dots, q$ ,  $q \rightarrow \infty$ .

Приведенные результаты могут быть обобщены на случай зависимых распределений случайных узлов, а также для случаев, когда вероятностный закон распределения узлов неизвестен и когда значения функции в узлах получены со случайными ошибками [4].

#### СПИСОК ЛИТЕРАТУРЫ

1. Теоретические основы и конструирование численных алгоритмов задач математической физики / Под ред. К. И. Бабенко. — М.: Наука, 1979.
2. Бахвалов Н.С. Об оптимальных оценках сходимости квадратурных процессов и методов интегрирования типа Монте-Карло на классах функций // Числ. методы решения дифференц. и интегр. уравнений и квадратурные формулы. — М.: Наука, 1964. — С. 5–63.
3. Корякин А.И. Рандомизированная оценка коэффициентов Фурье разложения функции в  $L_2$  // Докл. АН СССР. 1980. Т. 251, № 6. С. 1301–1305.
4. Корякин А.И., Ченцов Н.Н. Оценивание функций и их моментов по наблюдениям в случайных узлах. — М., 1985. — 24 с. — (Препр. / ИПМ; № 33).

## О ПРИМЕНЕНИИ МЕТОДА СЛУЧАЙНЫХ ИСПЫТАНИЙ (МЕТОДА МОНТЕ-КАРЛО) ДЛЯ РЕШЕНИЯ КИНЕТИЧЕСКОГО УРАВНЕНИЯ<sup>\*)</sup>

Обычно реакторы или другие мультиплицирующие системы рассчитывают, решая кинетическое уравнение. При этом значительную часть получаемой информации составляет всего лишь несколько числовых характеристик, таких, как эффективный коэффициент размножения, коэффициент размножения на быстрых нейтронах, эффективное время жизни нейтрона и т.п. В настоящей работе вычисление нужных характеристик системы будет сводиться к вычислению средних значений различных функционалов от нейтронных траекторий.

Если пренебрегать взаимодействием между нейтронами, то движение отдельного нейтрона в шестимерном фазовом пространстве является марковским случайным процессом (т.е. процессом без последствия). Соответствующее распределение  $P\{\cdot\}$  вероятностей в пространстве траекторий нейтрона полностью описывается положением в начальный момент времени и переходными вероятностями, которые в свою очередь определяются физическими свойствами окружающей среды: ее составом, сечениями ядерных реакций и законами деления и рассеяния. При таком подходе каждый функционал  $\varphi$  от траектории  $\xi(t)$  становится случайной величиной, а его среднее значение  $\varphi[\xi(\cdot)]$  — ее математическим ожиданием. Интеграл  $\int \varphi[x(\cdot)]P\{dx(\cdot)\}$ , изображающий это математическое ожидание, является интегралом по некоторой мере  $P$  в функциональном пространстве шестимерных функций  $x(t)$ , т.е. континуальным интегралом<sup>1)</sup>. Аналогичным образом можно рассматривать движение лавины нейтронов. В этом случае распределение вероятностей будет задаваться в пространстве “деревьев” нейтронных траекторий.

### Примеры

Пусть функция  $K(x, B)$  — математическое ожидание числа нейтронов, непосредственно порождаемых в области  $B$  одним нейтроном, который сам родился в точке  $x$  фазового пространства. Можно записать выражение

$$K(x, B) = \overline{\nu[\xi_x(\cdot)] \cdot \chi_B(\xi_x(\tau))}, \quad (1)$$

где  $\nu[\xi_x(\cdot)] = \nu(\xi_x(\tau))$  — случайное число нейтронов, образовавшихся после поглощения нейтрона при делении ядра-поглотителя (в случае вылета из системы

<sup>\*)</sup> Труды II международной конференции ООН по мирному использованию атомной энергии (Женева, 1958). Т. 2. — М.: Атомиздат. 1959. — С. 628–633 (совм. с И.М. Гельфандом, А.С. Фроловым, С.М. Фейнбергом).

<sup>1)</sup> В силу специфики движения нейтрона такой, вообще говоря, бесконечно кратный интеграл может быть представлен в виде бесконечной суммы кратных интегралов возрастающей кратности — по числу звеньев траектории нейтрона.

$\nu[\xi(\cdot)] = 0$ ;  $\tau = \tau[\xi(\cdot)]$  — случайный момент поглощения (или вылета);  $\chi_B(y) = 1$  при  $y \in B$ ;  $\chi_B(y) = 0$  при  $y \notin B$ . Среднее берется по всем возможным траекториям  $\xi_x(t)$  нейтрона, родившегося в момент  $t = 0$  в точке  $x$ . Далее, плотность распределения функционала  $\xi_x(\tau) \cdot \nu(\xi)$  будет плотностью рождения нейтронов, порождаемых источником, помещенным в точке  $x$ , а функция  $k(x) = \nu[\xi_x(\cdot)]$  является коэффициентом размножения для нейтронов этого источника. Усреднение функции  $k(x)$  по всем положениям рождения с весом, пропорциональным плотности рождения нейтронов, даст эффективный коэффициент размножения системы  $k_{\text{эф}}$ .

Часто требуется не только рассчитать среднее значение какого-либо функционала, но и все его распределение. Например, при расчете критической системы важно знать плотность рождения нейтронов. Выберем какую-либо систему ортонормированных функций  $h_k(z)$ . Коэффициенты  $a_k$  разложения по системе  $h_k(z)$  неизвестной плотности  $p(z)$  распределения функционала  $\varphi[\xi(\cdot)]$  можно записать как

$$a_k = \int h_k(z)p(z) dz = \int h_k\varphi[x(\cdot)] dP = \overline{h_k(\varphi[\xi(\cdot)])}.$$

При соответствующем выборе функций  $h_k(z)$  аппроксимирующие полиномы  $P_n(z) = \sum_{k=1}^n a_k h_k(z)$  дают достаточно хорошее приближение к неизвестному графику  $p(z)$ . Таким образом, приближенное вычисление распределения функционала  $\varphi$  можно свести к определению средних значений некоторых других функционалов  $h_k(\varphi)$ .

Для вычисления средних случайный процесс численно моделируется. По результатам случайных экспериментов конструируются ломаные, имитирующие траектории частиц. Единственное требование, предъявляемое к конструкции, — чтобы распределение вероятностей в пространстве функций  $x(t)$ , индуцируемое сконструированными случайными траекториями, совпадало с распределением вероятностей для моделируемого процесса. В качестве оценки искомого среднего значения  $\varphi[\xi(\cdot)]$  принимается среднее арифметическое значение  $\varphi$  от экспериментальных траекторий. Ошибка, которую дает такая своеобразная квадратурная формула, будет случайной, но для нее можно дать статистическую достоверную оценку. Обычно принимается, что ошибка не превышает  $3\sqrt{D(\varphi)/N}$ , где

$$D(\varphi) = \int \varphi^2 dP - \left( \int \varphi dP \right)^2$$

и  $N$  — число сконструированных траекторий.

Функционал  $\nu[\xi(\cdot)]$ , как и многие другие важные функционалы, заданные на траекториях нейтронов, принимает только целые значения. Для таких целочисленных функционалов отклонение экспериментального среднего от истинного значения велико, так что вычисление интеграла прямым моделированием требует слишком большого объема работы. Поэтому соответствующий континуальный интеграл приходится преобразовывать, приводя его к виду среднего значения другого функционала по траекториям некоторого фиктивного процесса.

Известен ряд способов преобразований, однако в каждой отдельной задаче целесообразно искать свою оптимальную модель. Этот поиск можно проводить

не только на основе теоретического анализа, но и по результатам небольших численных экспериментов. При этом качество модели оценивается количеством машинного времени, необходимым для достижения заданной точности. Такой подход, в частности, позволяет предоставить выбор оптимальной модели самой вычислительной машине.

Само построение траекторий удобно производить, разыгрывая последовательно историю частицы. Для построения используются бесконечные последовательности независимых случайных чисел, равномерно распределенных на отрезке  $[0, 1]$ . По этим числам, исходя из законов движения частицы, последовательно определяем элемент, с ядром которого столкнулась частица, тип столкновения, новую скорость и длину следующего пробега. В случае захвата или вылета частицы из системы просчет траектории заканчивается, и по новой последовательности начинается рассчитываться новая траектория. Таким образом, каждой бесконечной последовательности точек отрезка  $[0, 1]$  ставится в соответствие некоторая траектория. Тем самым задается отображение  $S$  счетномерного единичного гиперкуба  $H = [0, 1]^\infty$  на пространство траекторий, при котором лебегов объем  $V$  в  $H$  переходит в распределение вероятностей  $P$ , задающее моделируемый процесс. Каждый функционал  $\varphi[\xi]$  от траектории при этом переходит в функцию  $\tilde{\varphi}(y) = \varphi[Sy]$  от точки гиперкуба  $H$ .

Отметим одно важное свойство функций типа  $\tilde{\varphi}$ . Почти все траектории нейтронов состоят из конечного числа звеньев. Следовательно, для почти всех наборов аргументов функция  $\tilde{\varphi}$  будет зависеть только от конечного числа первых аргументов. В фиктивных процессах могут быть и бесконечные траектории, однако и здесь функции типа  $\tilde{\varphi}$  почти не зависят от далеких переменных. Так как интеграл по распределению  $P = VS^{-1}$  сводится в интеграл по объему  $V$ , то для вычисления средних нам достаточно научиться строить квадратурные формулы для функций типа  $\tilde{\varphi}$  на  $H$ .

По ряду причин при машинном счете неудобно использовать какие-либо внешние датчики случайных чисел. Поэтому вместо случайных чисел обычно используют так называемые псевдослучайные числа, генерируемые в самой машине.

Интересный метод получения последовательности псевдослучайных точек был предложен И.М. Коробовым. Для гладкой функции  $f(y)$  на  $H$ , зависящей от первых  $r$  переменных, он дает точную оценку:

$$\left| \frac{1}{N} \sum_{k=1}^N f(y_k) - \int_H f(y) dV(y) \right| = O\left(\frac{r-1}{\sqrt{N}}\right).$$

Здесь координаты  $y_k^{(s)}$  точек  $y_k$  задаются формулой  $y_k^{(s)} = \{g^{(2k+1)s}/p\}$ , где  $p = N$  — простое число вида  $p = 2p_1 + 1$ ;  $g$  — первообразный корень по модулю  $p$ ;  $\{z\}$  — дробная доля числа  $z$ . На другой возможный способ получения последовательности псевдослучайных точек указал И.И. Пятецкий-Шапиро. Если выбирать  $y_k^{(s)} = \{k\theta_s\}$ , то для каждой функции  $f(y)$  с конечной суммой  $Q$  абсолютных величин коэффициентов Фурье и почти всех наборов  $\theta_s$ ,  $0 < \theta_s < 1$ , имеет место более быстрая сходимость порядка  $O(Q \ln N/N)$ . Указанные последовательности обладают тем преимуществом, что обеспечивают также точную

сходимости:

$$\lim_{\substack{N \rightarrow \infty \\ k=1}} \frac{1}{N} \sum_{k=1}^N \varphi[sy_k] = \int \varphi[x(\cdot)] P\{dx(\cdot)\},$$

а не только сходимость по вероятности.

### Расчет реактора малых размеров

Для критического реактора (в стационарном состоянии) имеется следующий “закон сохранения”:

$$p(B) = \int \mathcal{K}(x, B) p(dx), \quad (2)$$

где через  $p(B)$  обозначено количество нейтронов, родившихся в области  $B$  за единицу времени, а функция  $\mathcal{K}(x, B)$  определена формулой (1). Введем оператор  $K$ , действующий в пространстве мер ограниченной вариации и задаваемый ядром  $\mathcal{K}(x, B)$  при помощи формулы

$$Kq(\cdot) = \int \mathcal{K}(x, \cdot) q(dx). \quad (3)$$

Тогда формула (2) выражает, что мера  $p(\cdot)$  есть собственная функция оператора  $K$ , отвечающая собственному числу  $k = 1$  (из физических соображений ясно, что  $k = 1$  — наибольшее собственное число оператора  $K$ ).

Предположим теперь, что ядро  $\mathcal{K}_c(x, B)$  и, стало быть, оператор  $K_c$  зависят от параметра  $c$ , причем при некотором значении  $c = c_0$  реактор становится критическим. При самых слабых физических допущениях операторы  $K_c$  оказываются положительными и, следовательно, имеют положительное собственное значение  $k_c$ , являющееся наибольшим по модулю собственным значением и отвечающее положительной собственной мере  $p_c(\cdot)$ . Другими словами, для  $K_c$  имеет место

$$K_c p_c(\cdot) = k_c p_c(\cdot) = \int \mathcal{K}_c(x, \cdot) p_c(dx), \quad (4)$$

причем при  $c = c_0$   $k_{c_0} = 1$ .

Таким образом, если мы сумеем находить собственные значения  $k_c$  при фиксированных  $c$ , то приближенное определение  $c_0$  не составит большого труда. Разыскивать собственную функцию оператора  $K_c$  довольно затруднительно, поэтому удобнее рассматривать сопряженный оператор  $K_c^*$ , имеющий тот же спектр и действующий в пространстве функций от точки по формуле

$$K_c^* f(\cdot) = \int \mathcal{K}_c(\cdot, dy) f(y). \quad (5)$$

В интересующих нас задачах ядро  $\mathcal{K}_c(x, B)$  оказывается кусочно гладкой функцией от  $x$ . Поэтому оператор  $K_c^*$  переводит непрерывные функции в гладкие; тем самым его собственные функции также оказываются кусочно гладкими.

При каждом фиксированном  $x$  интеграл  $\int \mathcal{K}_c(x, dy) f(y)$  имеет простой смысл:

$$g(x) = \int \mathcal{K}_c(x, dy) f(y) = \overline{\nu \cdot f(\xi_x(\tau))}, \quad (6)$$

где среднее берется по всем возможным траекториям нейтрона, начинающимся при  $t = 0$  в точке  $x$ . Эту величину легко вычислить методом Монте-Карло. Поэтому предлагается следующая процедура определения  $k_c$ . Выбирается  $m$  точек  $x_1, \dots, x_m$  и  $n$  функций  $f_1(x), \dots, f_n(x)$ , линейно независимых на множестве точек  $\{x_j\}_{j=1}^m$ ,  $n \leq m$ . Пусть известны значения  $g_i(x) = \int \mathcal{K}(x, dy) f_i(y)$  при всех  $x = x_j$ ,  $j = 1, \dots, m$ . Тогда мы можем каким-нибудь способом аппроксимировать  $g_i(x)$  функцией  $\tilde{g}_i(x) = \sum_{s=1}^n \alpha_s^{(i)} f_s(x)$  и тем самым аппроксимировать оператор  $K^*$  конечномерным оператором  $K_n^*$ , действующим в пространстве линейных комбинаций базисных функций  $f_i(x)$  по формуле  $K_n^* f_i(\cdot) = \sum_{s=1}^n \alpha_s^{(i)} f_s(\cdot)$ . При достаточно большом  $n$  наибольшее собственное число оператора  $K_n^*$  будет близко к собственному числу  $k$  оператора  $K^*$ .

При вычислении значений  $g_i(x_j)$  методом Монте-Карло нетрудно заодно приближенно вычислить их средние квадраты уклонения  $\delta_{ij}$ . Поэтому целесообразно аппроксимировать функции  $g_i(x)$  полиномами  $\tilde{g}_i(x)$  по методу наименьших квадратов. Наибольшее собственное число матрицы, изображающей оператор  $K_n^*$  в базисе из  $f_i(x)$ , вычисляется каким-либо прямым методом. Кроме того, рассчитывается собственный функционал оператора  $K_n^{**}$ , который может быть использован для приближенного определения средних значений каких-либо функций  $u(x)$  по собственной мере  $p_c(dx)$ , в частности для вычисления коэффициентов разложения плотности рождений  $p_c(dx)/dx$  по системе ортогональных функций.

Сведение задачи к расчету средних  $\nu \cdot f(\xi(\tau))$  удобно тем, что вычисление этих континуальных интегралов можно сильно ускорить, в первую очередь, за счет преобразования самого интеграла. В частности, значительный эффект дает моделирование движения частицы с фиктивной массой, при котором поглощение и вылет не разыгрываются, а учитываются при каждом столкновении весовым множителем. Величина  $\nu \cdot f(\xi(\tau))$  при этом заменяется на величину

$$\sum_k m_{k-1} \frac{\nu \cdot \Sigma_f(E_k, \xi(t_k))}{\Sigma_t \cdot (E_k, \xi(t_k))} f(\xi(t_k)), \quad (7)$$

где  $m_k$  — масса, остающаяся после  $k$ -го столкновения;  $\xi(t_k)$  — положение частицы в момент  $k$ -го столкновения;  $\Sigma_t$  и  $\Sigma_f$  — полное сечение и сечение деления.

Просчет траектории кончается, как только масса  $m_k$  становится меньше заранее заданной константы.

Описанная выше методика дает хорошие результаты в случаях, когда длина миграции нейтронов имеет тот же порядок, что и размеры реактора. Тогда собственная функция оператора  $K^*$  хорошо приближается линейной комбинацией небольшого числа базисных функций  $f_i$ .

Если пренебрегать зависимостью спектра скоростей нейтронов деления от скорости делящего нейтрона, то в уравнении (5) возможно разделить переменные, что сильно упрощает решение задачи.

### Расчет коэффициента размножения на быстрых нейтронах

Уточняя обычное определение, будем понимать под коэффициентом  $\epsilon$  размножения на быстрых нейтронах отношение числа нейтронов, замедлившихся до энергии  $E_0$  — порога деления  $U^{238}$ , к числу нейтронов, получившихся при де-

лении на тепловых нейтронах. Распределение  $p(B)$  количества нейтронов, порожденных в области  $B$  за единицу времени медленными нейтронами, предполагаем известным (например, из эксперимента).

Введем функцию  $\mathcal{L}(x, B)$  — математическое ожидание числа нейтронов, непосредственно порождаемых в области  $B$  нейтроном до того, как его энергия станет меньше энергии  $E_0$  при условии, что он сам родился в точке  $x$  фазового пространства. Введем также функцию  $l(x)$  — вероятность нейтрону, порожденному в точке  $x$ , замедлиться до энергии меньше пороговой.

Рассмотрим оператор  $L$ , задаваемый ядром  $\mathcal{L}(x, B)$ . Будем считать начальные нейтроны нейтронами нулевого поколения. Тогда количество нейтронов  $k$ -го поколения, рождающихся в области за единицу времени, задается мерой  $q_k(\cdot) = L^k p(\cdot)$ . Обозначая

$$\int f(x) q(dx) = (f, q),$$

можно записать коэффициент размножения в следующем виде:

$$\varepsilon = \sum_{k=0}^{\infty} \int l(x) q_k(dx) = \left( l, \sum_{k=0}^{\infty} L^k p \right). \quad (8)$$

Переходя к сопряженным операторам, имеем

$$\varepsilon = ([E - L^*]^{-1} l, p). \quad (9)$$

Для приближенных расчетов оператора  $L^*$  аппроксимируется конечномерным оператором  $L_n^*$  аналогично тому, как это было сделано при расчете реактора. Функция  $l(x)$  вычисляется методом Монте-Карло для нескольких значений аргументов, после чего также заменяется аппроксимирующей линейной комбинацией  $l_n(x)$  от базисных функций  $f_i(x)$ ,  $i = 1, \dots, n$ . В качестве оценки  $\varepsilon$  принимается величина

$$\varepsilon_n = ([E - L_n^*]^{-1} l_n, p), \quad (10)$$

где интегрирование по  $p$  производится обычным образом.

При вычислении значений  $l(x)$  и  $g_i(x) = L^* f_i(\cdot)|_x$  большой эффект дает моделирование движения частиц с фиктивной массой при котором поглощение частиц и падение их энергий ниже  $E_0$  не разыгрываются, а учитываются весовым множителем. Применимы и другие методы ускорения, в частности использование зависимых испытаний.

Следует отметить, что коэффициент  $\varepsilon$  можно вычислять прямым моделированием каскадного процесса на быстрых нейтронах. Расчет возникающих в таком процессе “деревьев” траекторий не связан с техническими трудностями, но требует большего объема работы, необходимой для достижения заданной точности.

Задача сильно упрощается, если считать, что спектры скоростей нейтронов деления для  $U^{238}$  и для  $U^{235}$  хотя и различны между собой, но не зависят от скорости делящего нейтрона.

## ИСПОЛЬЗОВАНИЕ ЗАВИСИМЫХ ИСПЫТАНИЙ В МЕТОДЕ МОНТЕ-КАРЛО ДЛЯ ПОЛУЧЕНИЯ ГЛАДКИХ КРИВЫХ <sup>\*)</sup>

В докладе обсуждается применение метода Монте-Карло для приближенного вычисления функциональных зависимостей.

Как известно, метод Монте-Карло был предложен для отыскания средних значений величин по статистическим испытаниям. Чтобы применить этот метод в конкретной задаче, приходится прежде всего интерпретировать искомую величину как среднее значение, т.е. как усреднение некоторой функции по какому-либо распределению вероятностей. Поэтому можно считать, что метод Монте-Карло есть некоторый специальный способ вычисления определенных интегралов. (Только этот последний термин нужно понимать достаточно широко — в смысле интеграла типа Лебега по произвольной мере в произвольном измеримом пространстве.)

Квадратурные формулы со случайными узлами, получающиеся по методу Монте-Карло, имеют ряд специфических особенностей. Самая неприятная из них — малая точность. Качество всякой квадратурной формулы характеризуется объемом вычислительной работы, которую надо проделать для достижения заданной точности. Если оценка остаточного члена формулы имеет вид  $CN^{-\alpha}$ , где  $N$  — число узлов, то за объем работы для достижения  $\varepsilon$ -точности мы примем величину

$$R = N\tau = \tau C^{1/\alpha} \left(\frac{1}{\varepsilon}\right)^{1/\alpha}, \quad (1)$$

где  $\tau$  — время вычисления одного значения подынтегральной функции. (Время подстановки вычисленных значений в формулу обычно также пропорционально  $N$  и может быть учтено в величине  $\tau$ .)

Ошибка квадратурной формулы типа Монте-Карло является случайной, причем с ростом  $N$  ее максимальное возможное значение не убывает, а остается постоянным. Поэтому пользуются не точной, а доверительной оценкой ошибки, которая имеет место с некоторой вероятностью  $1 - \delta$ . Обычно она имеет вид

$$\gamma \sqrt{\mathbf{D}f} N^{1/2}, \quad \text{где} \quad \mathbf{D}f = \overline{[f(\omega)]^2} - \overline{[f(\omega)]}^2;$$

$\gamma = \gamma(\delta)$  — постоянная, зависящая от уровня доверия  $1 - \delta$ . Эта сходимость очень медленная, и объем работы

$$R = \tau \mathbf{D}f \gamma^2 \left(\frac{1}{\varepsilon}\right)^2 \quad (2)$$

быстро растет с уменьшением  $\varepsilon$ . По этой причине часто приходится преобразовывать первоначальный интеграл так, чтобы в новой стохастической модели величина  $\tau \mathbf{D}f$  была возможно меньшей. Возникающие здесь вопросы неоднократно обсуждались в литературе, ряд соответствующих ссылок дан в [5] и [2].

<sup>\*)</sup> Труды VI Всесоюзного совещания по теории вероятностей и математической статистике (Вильнюс, 1960). — Вильнюс: ГИПНА, 1962. — С. 425–437 (совм. с А.С. Фроловым).

При использовании метода Монте-Карло для приближенного вычисления функций, заданных определенными интегралами, зависящими от параметра, приходится сталкиваться с новым обстоятельством. Понятие расстояния между функциями нельзя ввести универсальным образом. Поэтому способ, дающий удовлетворительную скорость сходимости в смысле какого-то расстояния, может оказаться непригодным из-за громадного объема работы, если требуется сходимость в смысле другого расстояния<sup>1)</sup>.

Рассмотрим простейший способ аппроксимации неизвестного графика. Пусть задана функция  $f(\omega, \alpha)$  случайного аргумента  $\omega$  и параметра  $\alpha$ . Искомую функцию

$$g(\alpha) = \mathbf{M}f(\omega; \alpha) = \int f(\omega; \alpha) P(d\omega)$$

вычисляем приближенно методом Монте-Карло при некоторых значениях параметра  $\alpha_1, \dots, \alpha_n$ , и по этим приближениям  $g^*(\alpha_i)$  проводим кривую. Этот способ неудобен тем, что при не слишком больших объемах выборок он не дает гладкого решения. Из-за статистического разброса точки  $g^*(\alpha_i)$  не лягут на гладкую кривую, хотя бы  $g(\alpha)$  была гладкой и  $f(\omega, \alpha)$  гладко зависела бы от  $\alpha$ .

Предлагается следующий прием. Пусть  $f(\omega; \alpha)$  — гладкая функция,  $\alpha, \omega_i$  — независимые испытания величины  $\omega$ . Тогда график  $\frac{1}{N} \sum_{i=1}^N f(\omega_i; \alpha)$  будет близок по форме к графику  $\mathbf{M}f(\omega; \alpha)$  даже при не слишком больших значениях  $N$ . Таким образом, если брать при всех значениях параметра одну и ту же последовательность наблюдений  $\omega$ , то можно быстро получить хорошее представление о графике  $g(\alpha)$ .

Дадим точную формулировку. Обозначим через  $\tilde{f}(\omega; \alpha)$  разность  $\tilde{f} = f(\omega; \alpha) - g(\alpha)$  и введем следующую характеристику гладкости  $f(\omega; \alpha)$ :

$$\left| \frac{\partial^k \tilde{f}(\omega; \alpha)}{\partial \alpha^k} \right| \leq G_k(\omega), \quad (\Gamma_k)$$

$$\int [G_k(\omega)]^p \mathbf{P}(d\omega) = C_k < \infty \quad (p \geq 2).$$

**ТЕОРЕМА 1.** Если функция  $f(\omega; \alpha)$  обладает характеристиками гладкости  $(\Gamma_k)$  ( $k = 0, \dots, n$ ), то случайная вектор-функция  $\vec{\varphi}_N$ ,

$$\vec{\varphi}_N = \sqrt{N} \left\{ \frac{1}{N} \sum_{i=1}^N f^{(k)}(\omega_i, \alpha) - g^{(k)}(\alpha) \right\}_{k=0}^{k=n-1},$$

сходится к некоторой случайной непрерывной гауссовской вектор-функции.

**СЛЕДСТВИЕ.** Если все условия  $\Gamma_0 - \Gamma_n$  выполнены, то уклонение эмпирической средней  $g_N^*(\alpha) = \frac{1}{N} \sum_{i=1}^N f(\omega_i; \alpha)$  и ее первых  $n - 1$  производных от истинной средней и ее производных имеет по вероятности порядок  $1/\sqrt{N}$ .

<sup>1)</sup>С этим обстоятельством, разумеется, приходится считаться и при вычислениях другими методами, но в методе Монте-Карло из-за медленной сходимости оно проявляется особенно сильно.

Доказательство теоремы. Заменяя  $f(\omega; \alpha)$  на  $\tilde{f}(\omega; \alpha)$ , имеем

$$\varphi_N^{(k)} = \frac{1}{\sqrt{N}} \sum_{i=1}^N \tilde{f}(\omega_i; \alpha).$$

Из условия  $(\Gamma_k)$  следует, что

$$\mathbf{D}\varphi_N^{(k)} \leq C_k.$$

Далее, оценив разность  $\varphi_N^{(k)}(\alpha'') - \varphi_N^{(k)}(\alpha')$  по формуле конечных приращений, можно удостовериться, что

$$\mathbf{M}|\varphi_N^{(k)}(\alpha'') - \varphi_N^{(k)}(\alpha')| \leq (\alpha'' - \alpha')^p C_{k+1} \quad (p > 1), \quad (3)$$

так что для  $\varphi_N^{(k)}$ ,  $k = 0, \dots, n-1$ , выполнен равномерно по  $N$  критерий непрерывности траекторий Колмогорова. Поэтому согласно [3] непрерывная случайная вектор-функция  $\vec{\varphi}_N(\alpha)$  слабо сходится к предельной случайной непрерывной гауссовской вектор-функции  $\vec{\varphi}_0(\alpha)$  и корреляционной функцией  $B_{jk}(\alpha', \alpha'')$  к средним  $m_k(\alpha)$ ,

$$B_{jk}(\alpha', \alpha'') = \mathbf{M}[f^{(j)}(\omega; \alpha') - g^{(j)}(\alpha')] [f^{(k)}(\omega; \alpha'') - g^{(k)}(\alpha'')], \\ m_k(\alpha) \equiv 0.$$

Таким образом, распределение непрерывного функционала — максимума модуля вектор-функции траектории — также слабо сходится к соответственному предельному распределению. В силу непрерывности траекторий предельного процесса они ограничены по вероятности.

Следовательно, отклонения эмпирических средних от теоретических, умноженные на  $\sqrt{N}$ , также ограничены по вероятности, а сами отклонения убывают по вероятности как  $1/\sqrt{N^2}$ .

Сразу же оговоримся, что предложенный выше способ выгоден далеко не всегда. Простейший пример:  $f(\omega; \alpha)$  не зависит от  $\alpha$ , т.е.  $g(\alpha) = \text{const}$ . Ясно, что, выбирая  $\omega_i$  независимыми при разных  $\alpha$  и проводя прямую по методу наименьших квадратов, мы получим большую точность при том же объеме работы. Вообще, когда заранее известно, что  $g(\alpha)$  предстоит выбирать из конечномерного множества функций, следует, по-видимому, применять методы оценок параметров. Наоборот, когда о вычисляемой функции заранее ничего не известно, применение одной и той же серии испытаний всегда оправдано. Опасно лишь, увидев, что точки легли на гладкую кривую, переоценить точность вычислений.

Предлагаемый способ особенно удобен, когда приходится вычислять производные  $g(\alpha)$ . Во многих интересных случаях вычисление  $\left. \frac{\partial f(\omega; \alpha)}{\partial \alpha} \right|_{\alpha=\alpha_0}$  при фиксированном  $\omega$  требует слишком большого объема работы, так что непосредственное вычисление методом Монте-Карло  $\left. \frac{\partial g}{\partial \alpha} \right|_{\alpha=\alpha_0}$  как среднего значения вели-

<sup>2)</sup>Идея оценки отклонения с помощью предельных теорем для случайных процессов была выдвинута А.А. Филипповой и Н.Н. Ченцовым и применена в [4].

чины  $\left. \frac{\partial f(\omega; \alpha)}{\partial \alpha} \right|_{\alpha=\alpha_0}$  невозможно. Если взять простейшую<sup>3)</sup> формулу и аппроксимировать  $g'(\alpha_0)$  разделенной разностью

$$\frac{g(\alpha_0 + h) - g(\alpha_0 - h)}{2h},$$

вычисляя значения  $g(\alpha_0 + h)$  и  $g(\alpha_0 - h)$  по двум независимым сериям испытаний, то в лучшем случае доверительная ошибка будет асимптотически равна

$$[9\gamma^2 \mathbf{D}(\alpha_0) g'''(\alpha_0)]^{1/3} \cdot (2N)^{-1/3} = O(N^{-1/3}),$$

где  $N$  — число испытаний в каждой серии, а

$$\mathbf{D}(\alpha_0) = \int [f(\omega; \alpha)]^2 \mathbf{P}(d\omega) - [g(\alpha_0)]^2.$$

Оптимальное значение  $h$  при этом равно  $(3\gamma)^{1/3} [g'''(\alpha_0)]^{-1/3} [\mathbf{D}(\alpha_0)]^{1/6} (2N)^{-1/6}$ , а объем работы, необходимый для достижения точности  $\varepsilon$ , равен

$$R(\varepsilon) = 2N\tau = \tau\gamma^2 [9\mathbf{D}(\alpha_0) g'''(\alpha_0)] \left(\frac{1}{\varepsilon}\right)^3. \quad (4)$$

При использовании одной и той же серии испытаний  $\omega$  шаг  $h$  можно брать сколько угодно малым (лишь бы не терялись при вычислении нужные значащие цифры). Поэтому случайная ошибка будет целиком определяться флуктуацией производной  $\partial\varphi_N^{(1)}/\partial\alpha$ , а систематическая ошибка, связанная с уклонением  $g(\alpha)$  от линейной функции  $g(\alpha_0) + (\alpha - \alpha_0)g'(\alpha_0)$ , отсутствует. Доверительная ошибка будет равна  $\gamma[\mathbf{D}_1(\alpha_0)]^{1/2} N^{-1/2}$ , где  $N$  — число испытаний в серии, а  $\mathbf{D}_1(\alpha_0) = \int [f'_\alpha(\omega; \alpha)]^2 \mathbf{P}(d\omega) - [g'(\alpha_0)]^2$ . Объем работы равен

$$R(\varepsilon) = 2N\tau = \tau\gamma^2 [2\mathbf{D}_1(\alpha_0)] \left(\frac{1}{\varepsilon}\right)^2, \quad (5)$$

т.е. по порядку существенно меньше. Время  $\tau$  вычисления одного значения функции  $f(\omega; \alpha)$  в обоих случаях одинаковое.

Отметим еще два способа приближенного вычисления функциональной зависимости  $g(\alpha)$ . Выберем густую сетку значений  $\alpha$  и сделаем при каждом значении несколько испытаний. Затем методом наименьших квадратов проведем плавную кривую. Для использования этого способа желательно заранее иметь некоторую информацию о  $g(\alpha)$ . Например, наличие у  $g(\alpha)$  излома необходимо учитывать заранее, а в способе одной серии излом изобразится автоматически.

Другой способ — искать разложение  $g(\alpha)$  в быстросходящийся ряд по удачно выбранной системе ортонормированных с весом  $p(\alpha)$  функций  $\varphi_k(\alpha)$ . Обозначив  $I = \int p(\alpha) d\alpha$ , можно записать коэффициенты  $a_k$  ряда  $g(\alpha) = \sum_{k=1}^{\infty} a_k \varphi_k(\alpha)$  как

$$a_k = \int g(\alpha) \varphi_k(\alpha) p(\alpha) d\alpha = \int_{\alpha} \int_{\omega} [I \varphi_k(\alpha) f(\omega; \alpha)] \mathbf{P}(d\omega) \frac{1}{I} p(\alpha) d\alpha. \quad (6)$$

<sup>3)</sup> Для функций  $g(\alpha)$ , обладающих большим запасом гладкости, можно применить более сложные, но зато и более точные формулы.

Последний интеграл можно истолковать как математическое ожидание функции  $I \varphi_k(\alpha) f(\omega; \alpha)$  двух независимых случайных аргументов  $\omega$  и  $\alpha$ , где  $\omega$  распределено по закону  $\mathbf{P}(d\omega)$ , а  $\alpha$  — по закону с плотностью  $p(\alpha)/I$ . Таким образом, коэффициенты  $a_k$  можно приближенно вычислять методом Монте-Карло, разыгрывая и аргумент  $\alpha$ , и аргумент  $\omega$ . В качестве приближения к  $g(\alpha)$  берется конечная сумма  $g_n^*(\alpha) = \sum_{k=1}^n a_k^* \varphi_k(\alpha)$ . Этот способ тоже требует предварительного знания свойств для того, чтобы правильно подобрать систему  $\varphi_k(\alpha)$ .

При решении многих конкретных задач подынтегральная функция  $f(\omega; \alpha)$  терпит разрыв как функция  $\alpha$  при фиксированном  $\omega$ . Может оказаться выгодным преобразовать интеграл  $\mathbf{M}f_1(\omega_1; \alpha) = \int f_1(\omega_1; \alpha) P_1(d\omega_1)$  к виду  $\mathbf{M}f_2(\omega_2; \alpha)$ , где  $f_2(\omega; \alpha)$  — уже непрерывная функция  $\alpha$ , хотя и более сложного строения. Замена бывает особенно выгодной, если удастся уменьшить “удельную работу” — коэффициент  $\tau \mathbf{D}f \gamma^2$ . Но даже если  $\tau_2 \mathbf{D}f_2 > \tau_1 \mathbf{D}f_1$ , общий объем работы для решения задачи может уменьшиться за счет того, что при гладком графике потребуется вычислять  $g(\alpha)$  с большой точностью для меньшего набора значений  $\alpha$ . Так бывает, например, при решении уравнения  $g(\alpha) = 0$ , когда ответом задачи является одно лишь число  $\alpha_0$  — корень уравнения  $g(\alpha) = 0$ .

Иногда для получения гладкого ответа достаточно разбить интеграл  $\mathbf{M}f(\omega; \alpha)$  на интегралы по подобластям гладкости и вычислять отдельно каждый интеграл по подобласти.

Рассмотрим важный пример со счетным числом подобластей гладкости. Пусть частица дискретным образом меняет свое положение в фазовом пространстве  $X$ . С вероятностью  $p_i(x; \alpha)$  происходит один из процессов  $A_1, \dots, A_m$ , переводящий частицу в новое положение, а с вероятностью  $q_i(x; \alpha)$  один из процессов  $\Pi_1, \dots, \Pi_n$  уничтожения частицы,  $\sum_{i=1}^m p_i + \sum_{j=1}^n q_j = 1$ . Каждый процесс  $A_i$  характеризуется вероятностью  $P_i(x, dy)$  перехода из точки  $x$  в окрестность  $dy$ , не зависящей от истории движения частицы. Таким образом, движение частицы до случайного момента  $\nu$  ее поглощения есть марковский процесс в фазовом пространстве  $X$  с дискретным временем.

Расширим фазовое пространство, взяв в качестве него произведение  $X \times B$ , где через  $B = \{\delta\}$  обозначена совокупность всех переходов  $A_i$  и всех поглощений  $\Pi_j$ . Такое “движение” с учетом всех имевших место актов также будет марковским процессом. Если для единообразия считать, что процессы  $\Pi_j$  переводят частицу в фиктивные состояния  $\tilde{x}$ , то траектория  $t$  движения будет описываться последовательностью

$$t = \{x_0, A^{(1)}, x_1, A^{(2)}, \dots, x_{j-1}, \Pi^{(\nu)}, \tilde{x}_\nu\}.$$

Так как процесс марковский, то естественно “разыгрывать” движение последовательно, от “столкновения” к “столкновению”. При этом приходится на каждом шаге разыгрывать, какой характер имеет “столкновение”. В качестве случая  $\omega$  возьмем последовательность независимых случайных величин  $\xi_i$ , равномерно распределенных на отрезке  $[0, 1]$  (или, что то же самое, случайную точку  $\omega$  единичного счетномерного гиперкуба с равномерным распределением). Функцией  $f(\omega; \alpha)$  будет какой-либо функционал от разыгранной по последовательности  $\omega$  траектории. Типичным интересным функционалом является аддитивный

функционал от траектории вида

$$f = \sum_{k=1}^{\nu} \varphi(x_{k-1}, \delta^{(k)}, x_k).$$

В каждом классе траекторий, задаваемом словом  $A^{(1)} \dots A^{(\nu-1)} \Pi^{(\nu)}$ , функционал  $f$  будет непрерывной функцией  $\alpha$ , но как только при изменении параметра  $\alpha$  моделируемая по  $\omega$  траектория перейдет в другой класс, функционал  $f$ , как правило, будет терпеть разрыв.

Мы предлагаем рассчитывать интеграл по каждому классу траекторий отдельно и затем вычислять их сумму<sup>4)</sup>. Разумеется, фактически надо вычислять лишь конечное число слагаемых, соответствующих высоковероятным словам. Количество их нетрудно приближенно оценить по следующей теореме.

**ТЕОРЕМА 2.** Пусть состояния  $A_1, \dots, A_m; \Pi_1, \dots, \Pi_n$  связаны в однородную цепь Маркова. Все состояния  $\Pi_j$  являются поглощающими, т.е.  $\mathbf{P}\{\Pi_j \rightarrow \Pi_j\} = 1$ . Вероятности перехода из остальных состояний

$$\mathbf{P}\{A_k \rightarrow A_i\} = p_i, \quad \mathbf{P}\{A_k \rightarrow \Pi_j\} = q_j$$

постоянны и не зависят от исходного состояния  $A_k$ .

Обозначим

$$P = \sum_{i=1}^m p_i, \quad H = - \sum_{i=1}^m \frac{p_i}{P} \log_2 \frac{p_i}{P}.$$

Тогда число высоковероятных цепочек  $N(\varepsilon)$  оценивается по порядку как

$$N(\varepsilon) \leq \left(\frac{1}{\varepsilon}\right)^{H/(-\log_2 P)} \quad \text{при } H \neq 0, \quad (7)$$

$$N(\varepsilon) \approx n \frac{\log_2 \varepsilon}{\log_2 P} \quad \text{при } H = 0. \quad (8)$$

**Доказательство.** Всякое блуждание вполне описывается последовательностью состояний до поглощения и поглотившим состоянием, т.е. некоторым словом из букв  $A_i$  и  $\Pi_j$ . Нетрудно подсчитать, что вероятность получить слово более, чем из  $s$  букв равна  $P^s$ . Выберем  $s$  так, чтобы  $P^s \leq \varepsilon$ . Для этого надо взять  $s \approx \log \varepsilon / \log P$ . При каждом  $k$  рассмотрим число  $N_k(\varepsilon)$  высоковероятных комбинаций букв  $A_i$ , появляющихся в схеме из  $k$  независимых испытаний с вероятностями  $p_i/P$ . Оно равно  $N_k(\varepsilon) \sim 2^{kH}$ . Вероятность появления остальных комбинаций букв  $A_i$  в схеме независимых испытаний не превосходит  $\varepsilon$ . Отсюда вероятность появления этих маловероятных комбинаций в цепи Маркова не превосходит  $\varepsilon P^k$ . Таким образом, число принятых во внимание комбинаций равно

$$\sum_{k=0}^{s-1} N_k(\varepsilon) \sim n \frac{1}{2^H - 1} 2^{sH} \sim \frac{n}{2^H - 1} \left(\frac{1}{\varepsilon}\right)^{H/(-\log_2 P)},$$

<sup>4)</sup>Метод и схема расчетов были впервые предложены А.С. Фроловым.

а вероятность всех отброшенных комбинаций равна

$$P^s + \varepsilon \sum_{k=1}^{s-1} P^k (1 - P) < 2\varepsilon.$$

Заменяя  $\varepsilon$  на  $\varepsilon/2$ , получаем по порядку опять величину

$$\left(\frac{1}{\varepsilon}\right)^{H/(-\log_2 P)}.$$

В случае вырождения  $H = 0$  обязательно  $P = p_k$ , а остальные  $p_i = 0$ . Таким образом, все слова имеют вид  $A_k \dots A_k \Pi_j$ ,  $j = 1, \dots, n$ . Число слов не более чем из  $s$  букв равно  $ns$ , т.е.  $N(\varepsilon) \approx n \log_2 \varepsilon / \log_2 P$ . Теорема доказана.

При вычислении интегралов отдельно по каждому классу траекторий в подынтегральное выражение входит “масса” частицы — якобиан преобразования от старых “координат” к новым в “точке”  $(x_0, A^{(1)}, x_1, \dots, x_{\nu-1}, \Pi^{(\nu)}, \tilde{x}_\nu)$ . Если разыгрывание траектории происходит последовательно, то розыгрыш варианта перехода отменяется и рекуррентно вычисляется “масса”  $\mu$  частицы:

$$\begin{aligned} \mu_0 &= 1, \\ \mu_k &= \mu_{k-1} p_i(x_{k-1}, \alpha), \quad \text{если } \delta^{(k)} = A_i, \\ \mu_\nu &= \mu_{\nu-1} q_j(x_{\nu-1}, \alpha), \quad \text{если } \delta^{(k)} = \Pi_j. \end{aligned} \quad (10)$$

Сам функционал приобретает вид

$$\tilde{f} = \mu_\nu \sum_{k=1}^{\nu} \varphi(x_{k-1}, \delta^{(k)}, x_k). \quad (11)$$

Особенно просто выглядит методика вычислений при  $H = 0$ . Здесь вместо дерева слов получается последовательность, и можно сразу просуммировать<sup>5)</sup> подынтегральные выражения:

$$\tilde{f} = \sum_{k=1}^{s(\omega)} \mu_{k-1} \left[ \varphi(x_{k-1}, A^{(k)}, x_k) p(x_{k-1}) + \sum_{j=1}^n \varphi(x_{k-1}, \Pi_j, \tilde{x}_{j,k}) q_j(x_{k-1}) \right], \quad (12)$$

где  $s(\omega)$  определяется из условия  $\mu_s < \varepsilon$ .

Разумеется, сведение линейного процесса к ветвящемуся выгодно далеко не всегда. Бывает, что ветвящийся процесс (например, при расчете каскада космических лучей) выгодно по разным соображениям моделировать линейным процессом. Здесь также приходится приписывать частицам массу, разыгрывая на каждом столкновении, за какой из образовавшихся частиц мы будем следить дальше, и умножая массу на число частиц, образовавшихся при столкновении. В частности, так поступил, по нашему совету, автор работы [6].

Кроме “чистых” тактик могут существовать “смешанные”. Выбор оптимальной модели можно вести с помощью экспериментов, как указано в [2].

<sup>5)</sup>Этот способ был описан нами ранее в [5] и [1].

Метод Монте-Карло можно применять не только для вычисления средних значений, но и для нахождения всего распределения случайной величины. Проще всего найти ее функцию распределения. Эмпирическая функция распределения всегда сходится к теоретической, и ее  $C$ -уклонение, по теореме Колмогорова [7], имеет порядок  $N^{-1/2}$ , где  $N$  — число проведенных независимых испытаний случайной величины. Значительно труднее найти плотность неизвестного распределения вероятностей. Обычно применяют метод гистограмм. Возможно (см. [1]), однако, применить более общие приемы.

**Поризм.** Пусть  $\xi$  — случайная величина (не обязательно числовая);  $\mu(dx)$  — мера в пространстве значений  $\xi$ ;  $p(x) = dP/d\mu$  — плотность распределения вероятностей  $\xi$  относительно меры  $\mu$ ;  $\varphi_k(x)$ ,  $k = 1, \dots, n$ , — система из  $n$  ортонормированных с весом  $r(x)$  функций,

$$\int \varphi_k(x) \varphi_l(x) r(x) \mu(dx) = \delta_{k,l}.$$

Пусть также

$$\|p(x)\|^2 = \int [p(x)]^2 r(x) \mu(dx) < \infty.$$

Тогда коэффициенты  $a_k$  среднеквадратичной аппроксимации

$$\pi(x) = \sum_{k=1}^n a_k \varphi_k(x)$$

неизвестной плотности  $p(x)$  могут быть истолкованы как

$$a_k = \mathbf{M}[\varphi_k(\xi) r(\xi)] \quad (13)$$

и приближенно сосчитаны методом Монте-Карло по независимым испытаниям  $\xi^{(1)}, \dots, \xi^{(N)}$  величины  $\xi$  как

$$a_k \approx \alpha_{k,N}^* = \frac{1}{N} \sum_{i=1}^N [\varphi_k(\xi^{(i)}) r(\xi^{(i)})], \quad (14)$$

а плотность  $p(x)$  приближена многочленом

$$\pi^*(x) = \sum_{k=1}^n \alpha_{k,N}^* \varphi_k(x).$$

Проверка соотношения (13) тривиальна:

$$a_k = \int p(x) \varphi_k(x) r(x) \mu(dx) = \int \varphi_k(x) r(x) d\mathbf{P} = \mathbf{M}[\varphi_k(\xi) r(\xi)].$$

Указанное математическое ожидание существует, так как

$$|a_k|^2 \leq \|p(x)\|^2 \|\varphi_k(x)\|^2 < \infty.$$

Обозначим для краткости  $\alpha_k = \varphi_k(\xi) r(\xi)$ . Поскольку  $\mathbf{M}\alpha_k = a_k$ , то по теореме Хинчина

$$\alpha_k^* = \frac{1}{N} \sum_{i=1}^N \alpha_k^{(i)} \Rightarrow a_k.$$

Чтобы дать оценки точности аппроксимации  $p(x)$  случайным многочленом  $\pi^*(x)$  в зависимости от числа  $N$  испытаний и “качества” базиса  $\varphi_k(x)$ , уточним понятие доверительной оценки.

ОПРЕДЕЛЕНИЕ. Число  $R = R(\delta)$  называется *отвечающей уровню доверия  $1-\delta$  доверительной оценкой (сверху)* случайной величины  $\rho$ , если  $\mathbf{P}\{\rho < R\} \geq 1 - \delta$ , или, что то же,

$$\mathbf{P}\{\rho \geq R\} \leq \delta.$$

Отметим, что если  $\mathbf{P}\{\rho > A\} > \Delta$ , то для любого  $\delta < \Delta$  и любой доверительной оценки  $R(\delta) > A$ .

ТЕОРЕМА 3. Пусть в обозначениях поризма задана последовательность весов  $r_n(x)$  и ортонормированных с весом  $r_n(x)$  систем  $\{\varphi_{k,n}(x)\}_{k=1}^n$  таких, что

$$\|p(x)\|_n^2 = \int [p(x)]^2 r_n(x) \mu(dx) \leq H < \infty. \quad (15)$$

А. Если

$$\|p(x) - \pi_n(x)\|^2 < C_1 n^{-\gamma}, \quad (16.1)$$

$$\int \frac{1}{n} \sum_{k=1}^n [\varphi_{k,n}(x) r_n(x)]^2 p(x) \mu(dx) < C_2, \quad (17.1)$$

то при выборе  $n \approx N^\beta$ ,  $\beta = 1/(1 + \gamma)$

$$\mathbf{P}\left\{\|p(x) - \pi^*(x)\| \geq \sqrt{\frac{C_1 + 2C_2}{\delta}} \cdot N^{-(1-\beta)/2}\right\} \leq \delta, \quad (19.1)$$

т.е.  $R(\delta) = \sqrt{(C_1 + 2C_2)/\delta} \cdot N^{-(1-\beta)/2}$  является доверительной оценкой уклонения  $\|p(x) - \pi^*(x)\|$ .

Б. Если

$$B_1 n^{-\gamma} < \|p(x) - \pi_n(x)\|^2, \quad (16.2)$$

$$4B_2 < \int \frac{1}{n} \sum_{k=1}^n [\varphi_{k,n}(x) r_n(x)]^2 p(x) \mu(dx), \quad (17.2)$$

$$\int \frac{1}{n^2} \sum_{k,l=1}^n [\varphi_{k,n}(x) r_n(x)]^2 [\varphi_{l,n}(x) r_n(x)]^2 p(x) \mu(dx) < C_3, \quad (18)$$

то, начиная с некоторого  $N$ , как бы ни подбиралось  $n$ , при всех достаточно малых  $\delta$  всякая доверительная оценка  $R(\delta) \geq BN^{-(1-\beta)/2}$ , где  $B$  — некоторая комбинация констант  $B_1$  и  $B_2$ , т.е.

$$\mathbf{P}\{\|p(x) - \pi^*(x)\| \geq BN^{-(1-\beta)/2}\} > \delta. \quad (19.2)$$

ЗАМЕЧАНИЕ. Условие (18) можно ослабить, заменив  $C_3$  на  $\tilde{C}_3 n^{1+\gamma}$ .

Доказательство. Рассмотрим разность  $p(x) - \pi^*(x)$ . Очевидно, она измерима по совокупности мер  $\mu(dx)$  и  $\mathbf{P}(d\omega)$ , так что ее можно интегрировать в любом порядке, не доказывая каждый раз возможности перестановки интегрирования. Уклонение  $\pi^*(x)$  от  $p(x)$  зависит от двух причин: от того, что  $\pi_n(x)$  не совпадает с  $p(x)$ , и от того, что  $\alpha_{k,n}^*$  уклоняются случайным образом от  $a_{k,n}$ . При квадратичных оценках оба эти уклонения складываются:

$$\begin{aligned} \|p(x) - \pi^*(x)\|^2 &= \int [p(x) - \pi^*(x)]^2 r_n(x) \mu(dx) = \\ &= \int [p(x) - \pi_n(x)]^2 r_n(x) \mu(dx) + \int [\pi_n(x) - \pi^*(x)]^2 r_n(x) \mu(dx) = \\ &= \|p(x) - \pi_n(x)\|^2 + \sum_{k=1}^n (\alpha_{k,n}^* - a_{k,n})^2, \end{aligned}$$

так как разность  $\pi_n(x) - \pi^*(x)$  лежит в пространстве линейных комбинаций  $\varphi_{k,n}(x)$ , а разность  $p(x) - \pi_n(x)$  перпендикулярна этому пространству. По условию теоремы  $\|p(x) - \pi_n(x)\|$  — неслучайная величина порядка  $n^{-\gamma}$ .

Оценим по порядку распределение величины  $\zeta^2 = \sum_{k=1}^n (\alpha_{k,n}^* - a_{k,n})^2$ :

$$\mathbf{M} \sum_{k=1}^n (\alpha_{k,n}^* - a_{k,n})^2 = \sum_{k=1}^n \mathbf{D}\alpha_{k,n} \frac{1}{N} \left\{ \sum_{k=1}^n \mathbf{M}[\alpha_{k,n}]^2 - \sum_{k=1}^n a_{k,n}^2 \right\}.$$

Из условий вытекает, что  $\mathbf{M}[\alpha_{k,n}]^2$  — величина порядка  $n$ . Из условия (15)  $\sum_{k=1}^n a_{k,n}^2 \leq \|p(x)\|_n^2 < H$ . Поэтому оба условия (17.1), (17.2) можно переписать в виде

$$\mathbf{M}\zeta^2 = \mathbf{M} \sum_{k=1}^n (\alpha_{k,n}^* - a_{k,n})^2 < C_2 \frac{n}{N}, \quad (17.3)$$

$$2B_2 \frac{n}{N} < \mathbf{M} \sum_{k=1}^n (\alpha_{k,n}^* - a_{k,n})^2 < C_2 \frac{n}{N} \quad \text{при} \quad n > \frac{H}{2B_2}. \quad (17.4)$$

Сложнее оценить  $\mathbf{M}\zeta^4$ . Применив ряд элементарных преобразований и неравенств, находим, что

$$\mathbf{M}\zeta^4 < \frac{1}{N} \left( \frac{n}{N} \right)^2 [C_3 + 4C_2 \sqrt{C_3} + 6(C_2)^2] + 3 \left( \frac{n}{N} \right)^2 (C_2)^2,$$

т.е. при  $C_4 = 9(C_2)^2 + 4C_2 \sqrt{C_3} + C_3$

$$\mathbf{M} \sum_{k,l=1}^n (\alpha_{k,n} - a_{k,n})^2 (\alpha_{l,n}^* - a_{l,n})^2 < C_4 \left( \frac{n}{N} \right)^2. \quad (18.1)$$

Перейдем к доказательству утверждения А. Имеем

$$\mathbf{M}\|p(x) - \pi^*(x)\|^2 = \|p(x) - \pi_n(x)\|^2 + \mathbf{M}\zeta^2 < C_1 \frac{1}{n^\gamma} + C_2 \frac{n}{N}.$$

Выбирая  $n = \lfloor N^\beta \rfloor + 1$ ,  $\beta = 1/(1 + \gamma)$ , получаем

$$\mathbf{M}\|p(x) - \pi^*(x)\|^2 < (C_1 + 2C_2)N^{-1+\beta}.$$

Отсюда, по неравенству Чебышева,

$$\mathbf{P}\left\{\|p(x) - \pi^*(x)\| \geq \sqrt{\frac{C_1 + 2C_2}{\delta}} N^{-(1-\beta)/2}\right\} < \delta.$$

Докажем теперь утверждение Б. Воспользуемся неравенством

$$\mathbf{P}\{\zeta^2 \geq \Theta M\zeta^2\} \geq \frac{[M\zeta^2]^2}{M\zeta^4} (1 - \Theta)^2. \quad (20)$$

Таким образом, с вероятностью, отличной от нуля,  $\zeta^2$  имеет тот же порядок, что и  $M\zeta^2$ .

Положив  $\Theta = 1/2$ , из (17.4) и (18.1) получаем

$$\mathbf{P}\left\{\zeta^2 \geq \frac{1}{2}M\zeta^2\right\} \geq \frac{4B_2^2}{C_4} \cdot \frac{1}{4}.$$

Следовательно, с вероятностью, большей, чем  $B_2^2/C_4$ ,

$$\begin{aligned} \|p(x) - \pi^*(x)\|^2 &= \|p(x) - \pi_n(x)\|^2 + \zeta^2 > \\ &> B_1 \frac{1}{n^\gamma} + B_2 \frac{n}{N} \geq [\gamma B_1 (B_2)^\gamma]^\beta \left(\frac{1+\gamma}{\gamma}\right) N^{-(1-\beta)}, \end{aligned}$$

причем знак равенства достигается лишь при  $n = [\gamma(B_1/B_2)N]^\beta$ . Поэтому при  $\delta < B_2^2/C_4$  имеем  $R(\delta) > BN^{-(1-\beta)/2}$ , где  $B = [\gamma B_1 (B_2)^\gamma]^\beta \sqrt{(1+\gamma)/\gamma}$ , так как

$$\mathbf{P}\{\|p(x) - \pi^*(x)\| \geq BN^{-(1-\beta)/2}\} > \delta.$$

**СЛЕДСТВИЕ 3.1.** <sup>6)</sup> Уклонение гистограммы случайной величины  $\xi$ ,  $a \leq \xi \leq b$ , от графика плотности при числе ступеней  $n \sim \sqrt[3]{N}$  в лучшем случае имеет по вероятности порядок  $N^{-1/3}$ .

**Доказательство.** Будем считать шаг гистограммы постоянным, т.е.  $h = (b-a)/n$ . В этом случае  $r(x) \equiv 1$ ,  $\mu(dx) = dx$ ,  $\varphi_{k,n}(x) = 0$  при  $x \notin I_k = [a + (k-1)h, a + kh]$ ,  $\varphi_{k,n}(x) = \sqrt{n/(b-a)}$  при  $x \in I_k$ . Условия (17) и (18) выполнены:

$$\begin{aligned} \int \frac{1}{n} \sum_{k=1}^n [\varphi_{k,n}(x)]^2 p(x) dx &= \frac{1}{n} \int_a^b \frac{n}{b-a} p(x) dx = \frac{1}{b-a}, \\ \int \frac{1}{n^2} \sum_{k,l=1}^n [\varphi_{k,n}(x)]^2 [\varphi_{l,n}(x)]^2 p(x) dx &= \frac{1}{n^2} \int_a^b \frac{n^2}{(b-a)^2} p(x) dx = \frac{1}{(b-a)^2}. \end{aligned}$$

<sup>6)</sup>Этот результат впервые был получен в [10].

Вычислим показатели  $\gamma$  и  $\beta$ . При  $x \in I_k$   $\pi(x) = \frac{1}{n} \int p(x) dx$ . Предполагая, что  $\frac{d^2 p}{dx^2}$  ограничена, получаем в главном члене

$$\int_a^b |p(x) - \pi_n(x)|^2 dx = \frac{h^2}{12} \int_a^b \left[ \frac{dp}{dx} \right]^2 dx + O(h^3) \sim n^{-2}.$$

Отсюда  $\beta = 1/3$ . По теореме 3 случайное уклонение  $\|p(x) - \pi^*(x)\|$  по порядку величины не менее  $N^{-1/3}$ . Эта оценка достигается при  $n \sim N^{1/3}$ , при другом выборе  $n$  она только хуже.

**СЛЕДСТВИЕ 3.2.** Уклонение гистограммы векторной случайной величины  $\vec{\xi} = \{\xi^{(m)}\}_{m=1}^s$ ,  $a_m \leq \xi^{(m)} \leq b_m$ , от графика плотности в лучшем случае при числе ступеней  $n \sim N^{s/(s+2)}$  имеет по вероятности порядок  $N^{-1/(2+s)}$ .

**Доказательство.** Будем считать шаги по каждой координате постоянными, т.е.  $h_m = (b_m - a_m)/n_m$ ,  $h = \max_m h_m$ ,  $n = \prod_{m=1}^s n_m$ ,

$$V = \prod_{m=1}^s (b_m - a_m), \quad h_0 = \left( \frac{V}{n} \right)^{1/s}, \quad h_m : h_0 = A_m.$$

Вес  $r(x) \equiv 1$ ,  $\mu(dx) = dx^{(1)} \dots dx^{(s)}$ . Введем интервалы

$$I(\vec{k}) = I(k_1, \dots, k_s) = \bigcap_{m=1}^s \{ \vec{x} : a_m + (k_m - 1)h_m \leq x^{(m)} \leq a_m + k_m h_m \};$$

$\varphi_{\vec{k},n}(\vec{x}) = 0$  при  $\vec{x} \notin I(\vec{k})$ ;  $\varphi_{\vec{k},n}(\vec{x}) = \sqrt{n/V}$  при  $\vec{x} \in I(\vec{k})$ . Условия (17) и (18) выполнены аналогично случаю  $s = 1$ :

$$\int \frac{1}{n} \sum_{\vec{k}} [\varphi_{\vec{k},n}(\vec{x})]^2 p(\vec{x}) d\vec{x} = \frac{1}{V},$$

$$\int \frac{1}{n} \sum_{\vec{k}, \vec{l}} [\varphi_{\vec{k},n}(\vec{x})]^2 [\varphi_{\vec{l},n}(\vec{x})]^2 p(\vec{x}) d\vec{x} = \frac{1}{V^2}.$$

Вычислим показатели. При  $\vec{x} \in I(\vec{k})$   $\pi_n(\vec{x}) = \frac{1}{h_1 \dots h_s} \int_I p(\vec{x}) d\vec{x}$ . Предполагая, что  $\frac{d^2 p(\vec{x})}{d\vec{x}^2}$  ограничена (т.е.  $\left| \frac{\partial^2 p}{\partial x_i \partial x_j} \right| \leq D < \infty$ ) при всех  $1 \leq i, j \leq s$  получаем в главном члене

$$\int |p(\vec{x}) - \pi_n(\vec{x})|^2 d\vec{x} = \frac{1}{12} (h_0)^2 \sum_{m=1}^s A_m^2 \int \left[ \frac{\partial p}{\partial x^{(m)}} \right]^2 d\vec{x} + O(h^3) \sim n^{-2/s}.$$

Отсюда, по теореме 3, случайное уклонение  $\|p(x) - \pi^*(x)\|$  по порядку величины не менее  $N^{-s/(s+2)}$ . Эта оценка достигается при  $n \sim N^{s/(s+2)}$ ; при другом выборе  $n$  она только хуже.

СЛЕДСТВИЕ 3.3. Пусть задана полная система ортонормированных с весом  $r(x)$  функций  $\varphi_k(x)$ . Если выполнены условия

$$\left| \int p(x) \varphi_k(x) r(x) \mu(dx) \right| < Ak^{-r}, \quad \int [\varphi_k(x) r(x)]^2 p(x) \mu(dx) < C,$$

то при выборе  $n \sim \sqrt[2r]{N}$  отклонение многочлена  $\pi^*(x)$  от  $p(x)$  в норме  $L_2(r)$  имеет по вероятности порядок  $N^{-1/2+1/(4r)}$ .

Доказательство. По равенству Парсеваля

$$\|p(x) - \pi_n(x)\|^2 = \sum_{k=n+1}^{\infty} a_k^2 < A(2r-1)n^{-2r+1},$$

откуда  $\gamma = 2r - 1$ ,  $C_1 = A(2r - 1)$ . Условие теоремы 3 также выполнено,  $C_2 = C$ .

Применяя теорему 3, получаем искомое.

Рассмотренные в настоящем докладе методы авторы разработали в связи с решением ряда задач вроде расчета критического размера реактора [1], расчета распространения  $\gamma$ -лучей в воздухе [9], расчета ветвящегося процесса [2] и др. Наша практика показала, что применение этих методов может быть полезным и дать существенный эффект.

#### СПИСОК ЛИТЕРАТУРЫ

1. Гельфанд И.М., Фейнберг С.М., Фролов А.С., Ченцов Н.Н. О применении метода случайных испытаний (метода Монте-Карло) для решения кинетического уравнения. Труды 2-й международной конференции ООН по применению атомной энергии в мирных целях (Женева, 1958, доклад 2142). Т. 2. — М., 1951. — С. 628–633.
2. Фролов А.С., Ченцов Н.Н. Некоторые вопросы численного конструирования случайных экспериментов // Методы программирования и решения задач на цифровых вычислительных машинах. — М., 1959. — С. 140–147.
3. Прохоров Ю.В. Сходимость случайных процессов и предельные теоремы теории вероятностей // Теория вероятн. и ее примен. 1956. Т. 1, вып. 2. С. 177–238.
4. Филиппова А.А. Теорема Мизеса о предельном поведении функционалов от эмпирических функций распределения и ее статистические применения: Дис. ... канд. физ.-мат. наук. — М.: МИ АН СССР, 1960.
5. Гельфанд И.М., Фролов А.С., Ченцов Н.Н. Вычисление континуальных интегралов методом Монте Карло // Изв. вузов. Сер. матем. 1958. Т. 5 (6). С. 32–45.
6. Голенко Д.И. Расчет характеристик некоторых статистических процессов методом Монте-Карло // Вычислительная математика. Вып. 5. — М.: Изд-во АН СССР, 1959. — С. 93–108.
7. Варфоломеев А.А., Голенко Д.И., Светлолобов И.А. Характеристики электромагнитных каскадов в фотоэмульсии с учетом влияния среды на процессы излучения // Докл. АН СССР. 1958. Т. 122, № 5. С. 785–787.
8. Колмогоров А.Н. Sulla determinazione empirico di una legge di distribuzione // Giorn. d. Att. 1933. V. 4. P. 83–91.
9. Лейпунский О.И., Стрелков А.С., Фролов А.С., Ченцов Н.Н. Распространение в воздухе гамма-квантов из мгновенного точечного источника // Атомная энергия. 1961. Т. 10, вып. 5. С. 493–500.
10. Смирнов Н.В. О построении доверительной области для плотности распределения случайной величины // Докл. АН СССР. 1950. Т. 74. С. 189–191.

## РЕШЕНИЕ ТРЕХ ТИПИЧНЫХ ЗАДАЧ ТЕОРИИ ПЕРЕНОСА МЕТОДОМ МОНТЕ-КАРЛО<sup>\*)</sup>

Статья посвящена решению задач физики переноса частиц методом Монте-Карло. Работа проводилась в период с 1955 г. по 1961 г. В процессе работы сложились некоторые общие способы и приемы решения таких задач. В статье делается попытка изложить методику решения задач физики переноса методом статистических испытаний с единой точки зрения. В общих чертах эта методика состоит в следующем:

- 1) искомые величины — характеристики исходного физического процесса — выражают через характеристики некоторого “производного” марковского случайного процесса;
- 2) траектории “производного” случайного процесса моделируют, и на них вычисляют значения некоторых специальных функционалов;
- 3) в результате статистической обработки числовой информации получают искомый ответ.

Процесс переноса обычно принято описывать интегро-дифференциальным уравнением Больцмана. Для целей стохастического моделирования более удобно пользоваться системой интегральных уравнений типа уравнения Пайерлса, описывающих процесс перехода от одного акта взаимодействия к другому. Соответствующая система уравнений (II) вводится в первом разделе статьи. Через решение этой системы непосредственно выражаются такие величины, как плотность числа рассеяний, поглощений и делений, плотность потока частиц и т.п. Более сложные характеристики процесса, например, период реактора или диффузионная длина, непосредственно через решение системы (II) не выражаются. Однако их можно выразить через средние значения некоторых специальных функционалов от траекторий частиц. Соответствующая мера в пространстве траекторий естественно задается ядрами интегральных операторов системы.

Во втором разделе рассмотрена задача о вычислении коэффициента размножения на быстрых нейтронах (см. также [1]). Искомые характеристики записываются через решение системы II и выражаются линейными формами от некоторых континуальных интегралов; коэффициенты форм определяются приближенно прямыми методами. Указываются приемы уменьшения объема вычислительной работы, приведен конкретный числовой результат.

Третий раздел статьи посвящен решению задачи о распространении гамма-излучения в воздухе (см. также [2]). Искомая интенсивность и ее средние по интервалам группировки записаны через плотность  $q$  решения системы (II). Описана модель случайного процесса и соответствующие функционалы от траекторий. Проведено сопоставление с результатами, полученными другими методами.

---

<sup>\*)</sup> Метод Монте-Карло в проблеме переноса излучений / Под ред. Г.И. Марчука. — М.: Атомиздат, 1967. — С. 25–52 (совм. с А.С. Фроловым).

В четвертом разделе описана методика расчета процессов рассеяния и захвата нейтронов в бороводородном сцинтиляторе (см. также [3]). Описана и обоснована вычислительная модель. Обсуждаются результаты расчетов, проведенных способом одной серии испытаний, что позволило получить характеристики сцинтиллятора в виде гладких кривых и найти полезные эмпирические зависимости.

### 1. Уравнения, описывающие процесс переноса

**Математическая модель процесса переноса.** Постулируются следующие свойства математической модели процесса переноса частиц:

- процесс движения частицы является марковским случайным процессом, однородным во времени;
- движения частиц в шестимерном фазовом пространстве  $\Omega$  координат  $\mathbf{x}$  и скоростей  $\mathbf{v}$  независимы друг от друга и имеют одинаковые вероятностные распределения;
- траектория  $\xi(t)$  отдельной частицы вполне определяется точками рождения, столкновения и гибели, а также скоростями, с которыми частица двигалась от столкновения к столкновению;
- соответствующие распределения вероятностей  $P\{\cdot\}$  в пространстве траекторий (деревьев траекторий) полностью описываются положениями в начальный момент  $t = 0$  и переходными вероятностями в последующие моменты времени.

Это допустимо, если в рассматриваемых физических системах:

- концентрация частиц в веществе мала и мала вероятность их взаимодействия, так что движения отдельных частиц не зависят друг от друга;
- вещество находится в стационарном состоянии;
- вероятность всех элементарных актов взаимодействия частицы с веществом определяется природой взаимодействующих частиц и не зависит от предыстории рассматриваемой частицы.

Таким образом, сведения о процессе переноса в рассматриваемой модели могут быть получены из статистики движения отдельной частицы (или статистики деревьев траекторий данной частицы и ее потомков).

Аналогично может быть рассмотрен процесс переноса частиц нескольких сортов.

**Стохастическое дифференциальное описание процесса.** В сделанных предположениях случайное движение частицы описывается следующим образом (см., например, работы [4] и [5]). Частица движется прямолинейно и равномерно до тех пор, пока не столкнется с ядром вещества. Вероятность столкновения на пути  $ds$  равна  $\Sigma(\mathbf{x}, |\mathbf{v}|) ds$ , где  $\Sigma(\mathbf{x}, |\mathbf{v}|)$  — макроскопическое сечение. Вероятность поглощения при этом —  $\Sigma_{\text{погл}}/\Sigma$ , вероятность рассеяния —  $\Sigma_{\text{рас}}/\Sigma$ , вероятность деления (образования нескольких частиц) —  $\Sigma_{\text{дел}}/\Sigma$ . Плотность вероятности изменить при рассеянии скорость  $\mathbf{v}'$  на скорость  $\mathbf{v}$  равна  $K(\mathbf{v}', \mathbf{v}, \mathbf{x})$  (индикатриса рассеяния). Аналогично, усредненная индикатриса деления (на одну частицу) равна  $\Lambda(\mathbf{v}', \mathbf{v}, \mathbf{x})$ , а  $\bar{\nu}(\mathbf{x}, \mathbf{v})$  есть среднее число новых частиц, образующихся в точке  $\mathbf{x}$  при делении частицы, имевшей скорость  $\mathbf{v}'$ . Введем также функцию источников  $\varphi(t, \mathbf{x}, \mathbf{v})$  — плотность числа частиц, испускаемых источником.

**Интегро-дифференциальное и интегральное уравнения, описывающие процесс переноса.** Для плотности  $u(t, \mathbf{x}, \mathbf{v})$  числа частиц, в момент  $t$  проходящих через точку  $\mathbf{x}$  со скоростью  $\mathbf{v}$ , можно написать известное интегро-дифференциальное уравнение переноса

$$\begin{aligned} \frac{\partial u(t, \mathbf{x}, \mathbf{v})}{\partial t} + (\mathbf{v} \operatorname{grad} u(t, \mathbf{x}, \mathbf{v})) &= \varphi(t, \mathbf{x}, \mathbf{v}) - |\mathbf{v}| \Sigma(\mathbf{x}, |\mathbf{v}|) u(t, \mathbf{x}, \mathbf{v}) + \\ &+ \int_{(\mathbf{v}')} |\mathbf{v}'| \Sigma_{\text{pac}}(\mathbf{x}, \mathbf{v}') K(\mathbf{v}', \mathbf{v}, \mathbf{x}) u(t, \mathbf{x}, \mathbf{v}') d\mathbf{v}' + \\ &+ \int_{(\mathbf{v}')} \bar{\nu}(\mathbf{x}, \mathbf{v}') |\mathbf{v}'| \Sigma_{\text{дел}}(\mathbf{x}, \mathbf{v}') \Lambda(\mathbf{v}', \mathbf{v}, \mathbf{x}) u(t, \mathbf{x}, \mathbf{v}') d\mathbf{v}'. \end{aligned} \quad (1.1)$$

Это уравнение справедливо и для плотности вероятности случайной частице пройти в момент  $t$  через точку  $(\mathbf{x}, \mathbf{v})$ .

Отметим следующее важное обстоятельство. Движение частицы состоит из последовательных повторений движений двух типов: изменения пространственных координат  $\mathbf{x}$  и изменения скорости  $\mathbf{v}$ . Обозначим  $q(t, \mathbf{x}, \mathbf{v})$  — плотность числа частиц, вступивших в точке  $(\mathbf{x}, \mathbf{v})$  в момент  $t$  во взаимодействие с веществом:

$$q(t, \mathbf{x}, \mathbf{v}) = |\mathbf{v}| \Sigma(\mathbf{x}, |\mathbf{v}|) u(t, \mathbf{x}, \mathbf{v}); \quad (1.2)$$

$p(t, \mathbf{x}, \mathbf{v})$  — плотность числа частиц, рассеявшихся в момент  $t$  в точке  $(\mathbf{x}, \mathbf{v})$ ;  $r(t, \mathbf{x}, \mathbf{v})$  — плотность числа частиц, возникших в момент  $t$  в точке  $(\mathbf{x}, \mathbf{v})$ . Эти функции и функция источников связаны между собой тремя интегральными соотношениями, которые при выполнении некоторых условий регулярности<sup>1)</sup> вытекают из уравнения переноса (1.1). Эти соотношения нам понадобятся только для стационарных состояний физической системы, когда функции  $p, q, r$  и  $\varphi$  не зависят от времени. В этом случае они переходят в систему трех интегральных уравнений

$$\text{(II)} \quad \begin{cases} q = A(p + r + \varphi), & (1.3) \\ p = Bq, & (1.4) \\ r = Cq, & (1.5) \end{cases}$$

с интегральными операторами  $A, B$  и  $C$ :

$$A[f(\mathbf{x})] = \int_{(\mathbf{x}')} R_1(\mathbf{x}', \mathbf{x}, \mathbf{v}) f(\mathbf{x}') d\mathbf{x}' = \Sigma(\mathbf{x}, |\mathbf{v}|) \int_0^\infty \rho(\mathbf{x}'(s); \mathbf{x}, |\mathbf{v}|) f(\mathbf{x}'(s)) ds, \quad (1.6)$$

где  $\mathbf{x}'(s) = \mathbf{x} - \frac{\mathbf{v}}{|\mathbf{v}|} s$  и

$$\rho(\mathbf{x}'(s); \mathbf{x}, |\mathbf{v}|) = \exp \left\{ - \int_0^s \Sigma(\mathbf{x}'(\theta), |\mathbf{v}|) d\theta \right\};$$

$$B[f(\mathbf{v})] = \int_{(\mathbf{v}')} R_2(\mathbf{v}', \mathbf{v}, \mathbf{x}) f(\mathbf{v}') d\mathbf{v}', \quad (1.7)$$

<sup>1)</sup> Например, условия, что с вероятностью 1 за конечное время история частицы состоит из конечного числа рассеяний и делений и т. п. (ср. [6]).

где

$$R_2(\mathbf{v}', \mathbf{v}, \mathbf{x}) = \frac{\Sigma_{\text{pac}}(\mathbf{x}, \mathbf{v}')}{\Sigma(\mathbf{x}, |\mathbf{v}'|)} K(\mathbf{v}', \mathbf{v}, \mathbf{x}),$$

$$C[f(\mathbf{v})] = \int_{(\mathbf{v}')} R_3(\mathbf{v}', \mathbf{v}, \mathbf{x}) f(\mathbf{v}') d\mathbf{v}', \quad (1.8)$$

где

$$R_3(\mathbf{v}', \mathbf{v}, \mathbf{x}) = \bar{\nu}(\mathbf{x}, \mathbf{v}') \frac{\Sigma_{\text{дел}}(\mathbf{x}, \mathbf{v}')}{\Sigma(\mathbf{x}, |\mathbf{v}'|)} \Lambda(\mathbf{v}', \mathbf{v}, \mathbf{x}).$$

Введенные выше индикатрисы рассеяния и деления  $K$  и  $\Lambda$  имеют различные виды в зависимости от типа переносимых частиц и состава вещества, в котором происходит перенос.

Вычисление распределений частиц проводится по следующему принципу: методом случайных испытаний решается система (II) или уравнения, ей эквивалентные. Очевидны следующие преобразования системы: исключив  $q$ , можно получить одно уравнение<sup>2)</sup> для  $p + r$

$$p(\mathbf{x}, \mathbf{v}) + r(\mathbf{x}, \mathbf{v}) = \int_{(\mathbf{v}')} \int_{(\mathbf{x}')} [R_2(\mathbf{v}', \mathbf{v}, \mathbf{x}) + R_3(\mathbf{v}', \mathbf{v}, \mathbf{x})] R_1(\mathbf{x}', \mathbf{x}, \mathbf{v}') \times \\ \times [p(\mathbf{x}', \mathbf{v}') + r(\mathbf{x}', \mathbf{v}') + \varphi(\mathbf{x}', \mathbf{v}')] d\mathbf{v}' d\mathbf{x}'. \quad (1.9)$$

Найдя  $p + r$ , по соотношению (1.3) досчитать  $q$  и, используя соотношения (1.4) и (1.5), по  $q$  рассчитать  $p$  и  $r$  по отдельности. Наоборот, можно исключить  $p$  и  $r$  и получить одно уравнение для  $q$ . Более сложная схема применена в расчете, описанном в работе [1], когда из системы исключается  $p$  и  $q$  и остается уравнение для  $r$ .

Если стационарный режим поддерживается источниками, решение системы интегральных уравнений (II) второго рода может быть получено из разложения в ряд<sup>3)</sup>: последовательно вычисляются

$$p_0 + r_0 = \varphi,$$

$$q_n(\mathbf{x}, \mathbf{v}) = \int_{(\mathbf{x}')} R_1(p_{n-1} + r_{n-1}) d\mathbf{x}', \quad (1.10)$$

$$p_n(\mathbf{x}, \mathbf{v}) = \int_{(\mathbf{v}')} R_2 q_n d\mathbf{v}'; \quad r_n(\mathbf{x}, \mathbf{v}) = \int_{(\mathbf{v}')} R_3 q_n d\mathbf{v}'.$$

Тогда

$$p = \sum_{n=1}^{\infty} p_n, \quad q = \sum_{n=1}^{\infty} q_n, \quad r = \sum_{n=1}^{\infty} r_n.$$

<sup>2)</sup> Можно было бы с самого начала ограничиться системой для  $q$  и  $p + r$ , рассматривая только два типа актов: изменение положения и изменение скорости. Наоборот, можно различать два типа рассеяния: упругое и неупругое. Вопросы изменения схемы моделирования, возникающие при подобном "склеивании" и "расщеплении" ядер операторов, обсуждаются в третьем разделе.

<sup>3)</sup> Обозначив тройку функций  $(q, p, r) = \mathbf{p}$  и линейный оператор, соответствующий правым частям системы (II), символом  $\Theta$ , получаем уравнение  $\mathbf{p} = \Theta \mathbf{p} + \tilde{\varphi}$  в операторной форме, где  $\tilde{\varphi} = (\varphi, 0, 0)$ . Его решение  $\mathbf{p} = (I - \Theta)^{-1} \tilde{\varphi} = \tilde{\varphi} + \Theta[\tilde{\varphi}] + \Theta^2[\tilde{\varphi}] + \dots$

Аналогичным образом можно решать системы интегральных уравнений, эквивалентных системе (П). Искомое решение выписывается в сходной форме.

Описанный выше подход пригоден не только для стационарных режимов физических систем, но и для некоторых стационарно меняющихся режимов, в частности для определения периода реактора  $T$  [5]. Если  $u(t, \mathbf{x}, \mathbf{v}) = e^{\pm t/T} u(\mathbf{x}, \mathbf{v})$ , то плотность  $u(\mathbf{x}, \mathbf{v})$  удовлетворяет стационарному уравнению переноса (1.1) при условии  $\partial u / \partial t = 0$  с добавочным сечением поглощения  $\pm 1 / (T|\mathbf{v}|)$ .

Всюду в нашей работе искомые величины будут выражаться через характеристики некоторого производного марковского процесса с точечным мгновенным источником.

## 2. Расчет коэффициента размножения на быстрых нейтронах

Рассматривалась физическая система, состоящая из  $U^{238}$  и замедлителя, с известным распределением источников высокоэнергичных нейтронов. Поток нейтронов из источника вызывает в уране затухающую цепную реакцию.

Требовалось оценить коэффициент усиления первоначального потока. Для этого подсчитывалось среднее число  $\varepsilon$  нейтронов, возникающих в цепной реакции и замедлившихся до энергии ниже  $E_0$  — пороговой энергии деления  $U^{238}$ .

Подобная задача представляет интерес при расчете систем из естественного урана. В этих системах подавляющая часть нейтронов деления образуется при делении  $U^{235}$  тепловыми нейтронами. Делением  $U^{235}$  в надтепловой области в силу малого его содержания можно пренебречь.

При энергиях больше  $E_0$   $U^{238}$  действует не только как поглотитель, но и как размножитель, грубо говоря, увеличивающий коэффициент размножения на  $U^{235}$ . Поэтому  $\varepsilon$  было условно названо коэффициентом размножения на быстрых нейтронах. Система интегральных уравнений переноса имеет самый общий вид:

$$q = A(p + r + \varphi), \quad p = Bq, \quad r = Cq. \quad (2.1)$$

Пусть  $\chi$  — характеристическая функция области энергии  $E \leq E_0$ , т.е.

$$\chi(\mathbf{x}, \mathbf{v}) = \begin{cases} 0, & E > E_0, \\ 1, & E \leq E_0, \end{cases}$$

и пусть  $J$  — оператор умножения на функцию  $\chi^+ = 1 - \chi$  — характеристическую функцию области энергии  $E > E_0$ . Определенный выше коэффициент размножения на быстрых нейтронах  $\varepsilon$  в этих обозначениях равен

$$\varepsilon = (\chi, r + \varphi + BJq), \quad (2.2)$$

если считать интенсивность источника единичной,  $(1, \varphi) = 1$ . Для расчетов выражение (2.2) полезно преобразовать. Обозначим  $l(\mathbf{x}, \mathbf{v})$  — вероятность замедления нейтрона до энергии меньше пороговой. Введем также оператор  $L$  с ядром  $L(\mathbf{x}', \mathbf{v}', \mathbf{x}, \mathbf{v})$  — плотностью математического ожидания числа нейтронов, непосредственно порождаемых в окрестности точки  $(\mathbf{x}, \mathbf{v})$  нейтроном до того, как

его энергия станет меньше энергии  $E_0$ , при условии, что он сам родился в точке  $(\mathbf{x}', \mathbf{v}')$  фазового пространства. Тогда

$$\varepsilon = \left( l, \sum_{k=0}^{\infty} L^k \varphi \right). \quad (2.3)$$

Выведем выражение (2.3) из формулы (2.2). Идемпотентный оператор  $J$  обладает следующими свойствами:

$$J^2 = J, \quad J = J^*, \quad C = CJ, \quad AJ = JA, \quad JB = JBJ. \quad (2.4)$$

Исключив из системы (2.1) функцию  $p$ , можно выразить плотность  $q$  через  $r + \varphi$ :  $q = [I - AB]^{-1}(r + \varphi)$ . Пользуясь свойствами оператора  $J$ , получаем

$$\begin{aligned} r + \varphi + VJq &= [I + VJ(I - AB)^{-1}A](r + \varphi) = \\ &= [I + VJ(I + (AJB) + \dots + (AJB)^n + \dots)AJ](r + \varphi). \end{aligned}$$

Обозначим

$$\Lambda = [I + (AJB) + \dots + (AJB)^n + \dots]AJ \quad (2.5)$$

и перейдем в скалярном произведении (2.2) к сопряженным операторам

$$\varepsilon = ([I + \Lambda^*JB]^* \chi, r + \varphi).$$

Первая из функций в этом скалярном произведении и есть  $l(\mathbf{x}, \mathbf{v})$ :

$$l(\mathbf{x}, \mathbf{v}) = [I + \Lambda^*JB]^* \chi. \quad (2.6)$$

Теперь выразим  $r + \varphi$  через  $\varphi$ :

$$r + \varphi = Cq + \varphi = C[I - AB]^{-1}A(r + \varphi) + \varphi.$$

Так как  $C = CJ$ , то можно записать

$$[I - C\Lambda](r + \varphi) = \varphi. \quad (2.7)$$

Нетрудно видеть, что  $C\Lambda$  совпадает с введенным ранее оператором  $L$ :  $L = C \cdot \Lambda$ . Тем самым вывод (2.3) завершен.

Чтобы избежать моделирования ветвящегося процесса, используем следующий прием. Выберем систему базисных функций  $f_i(\mathbf{x}, \mathbf{v})$ ,  $i = 1, \dots, n$ . Методом случайных испытаний вычислим функцию  $l(\mathbf{x}, \mathbf{v})$  для выбранных значений аргумента, после чего заменим ее аппроксимирующей линейной комбинацией базисных функций  $l_n = \sum_{i=1}^n \beta_i f_i$ . Аналогично аппроксимируем оператор  $L^*$  конечномерным оператором  $L_n^*$ , действующим в пространстве линейных комбинаций базисных функций  $f_i$ . Для каждой функции  $f_i$  вычислим методом случайных испытаний при избранных значениях аргумента  $(\mathbf{x}, \mathbf{v})$  значения  $L_n^* f_i$ . По этим значениям проводим аппроксимирующую линейную комбинацию базисных функций

$$\tilde{f}_i = \sum_{s=1}^n \alpha_s^{(i)} f_s.$$

В качестве оценки  $\varepsilon$  примем величину

$$\varepsilon_n = ([I - L_n^*]^{-1} l_n, \varphi), \quad (2.8)$$

где интегрирование по  $\varphi$  производится любым способом. Задача сильно упрощается, если считать, что спектры скоростей нейтронов деления для  $U^{238}$  и  $U^{235}$  хотя и различны между собой, не зависят от скорости делящего нейтрона. В этом случае

$$L(\mathbf{x}', \mathbf{v}', \mathbf{x}, \mathbf{v}) = L'(\mathbf{x}', \mathbf{v}', \mathbf{x}) \psi(\mathbf{v}), \quad \varphi(\mathbf{x}, \mathbf{v}) = \tilde{\varphi}(\mathbf{x}) \varphi(\mathbf{v}).$$

Тогда интегрирование по  $\mathbf{v}$  можно провести методом случайных испытаний, подсчитав

$$\begin{aligned} l_0(\mathbf{x}) &= \int l(\mathbf{x}, \mathbf{v}) \varphi(\mathbf{v}) d\mathbf{v}, \\ l_1(\mathbf{x}) &= \int l(\mathbf{x}, \mathbf{v}) \psi(\mathbf{v}) d\mathbf{v}, \\ L_0(\mathbf{x}', \mathbf{x}) &= \int L'(\mathbf{x}', \mathbf{v}', \mathbf{x}) \varphi(\mathbf{v}') d\mathbf{v}', \\ L_1(\mathbf{x}', \mathbf{x}) &= \int L'(\mathbf{x}', \mathbf{v}', \mathbf{x}) \psi(\mathbf{v}') d\mathbf{v}'. \end{aligned} \quad (2.9)$$

Далее,

$$\begin{aligned} \varepsilon &= (l, \varphi) + \sum_{k=1}^{\infty} (l, L^k \varphi), \\ \varepsilon &= \int l_0(\mathbf{x}) \cdot \tilde{\varphi}(\mathbf{x}) d\mathbf{x} + \sum_{k=1}^{\infty} l_1(\mathbf{x}) \cdot [L_1^{k-1} \cdot L_0 \tilde{\varphi}(\cdot)]_{\mathbf{x}} d\mathbf{x}. \end{aligned} \quad (2.10)$$

В случае однородного пространства формула (2.10) предельно упрощается:

$$\begin{aligned} \varepsilon &= l_0 + l_1 \sum_{k=1}^{\infty} L_1^{k-1} L_0, \\ \varepsilon &= l_0 + \frac{L_0 l_1}{1 - L_1}. \end{aligned} \quad (2.11)$$

Здесь  $l_0$  равно средней по спектру вероятности замедления нейтрона источника до энергии меньше пороговой;  $l_1$  — полная (по всему спектру деления) вероятность замедления нейтрона, родившегося при делении  $U^{238}$ , до энергии меньше пороговой;  $L_0$  — усредненное по спектру источника математическое ожидание числа нейтронов, возникающих при делении ядра  $U^{238}$  нейтроном источника;  $L_1$  — математическое ожидание числа нейтронов, возникающих при делении ядра  $U^{238}$  нейтроном деления.

Как видно из формул (2.9) для  $l$  и  $L$ , при их вычислениях можно моделировать исходный процесс в одном поколении, заменяя при этом розыгрыш акта деления дополнительным поглощением. По формуле (2.6)

$$l_0 = \tilde{l}_0 + \sum_{n=0}^{\infty} \int A^* (B^* J A^*)^n B^* \chi \varphi(\mathbf{v}) d\mathbf{v},$$

где  $\tilde{l}_0 = \int_0^{v(E_0)} \varphi(\mathbf{v}) d\mathbf{v}$ ,  $B^* \chi = \pi(E)$  — вероятность замедления от энергии  $E$  до энергии ниже пороговой. Столкновывая ядро  $A^*$  как ядро оператора переноса по пространству, ядро оператора  $B^*$  — как произведение вероятности рассеяния на индикатрису рассеяния,  $J$  — как фактор обрыва траектории при снижении энергии частицы до энергии ниже пороговой, можно моделировать движение частицы с розыгрышем значения начальной энергии по индикатрисе деления  $\varphi(v)$  до момента ее поглощения или уменьшения ее энергии ниже пороговой. На каждом положении вычисляется  $\pi(E_{n-1})$ , и вдоль траектории  $\xi$  подсчитываются

$$\begin{aligned} l_0(\xi) &= \tilde{l}_0 + \sum_{n=1}^{\infty} \pi(E_{n-1}), \\ l_1(\xi) &= \tilde{l}_1 + \frac{\psi(v_0)}{\varphi(v_0)} \sum_{n=1}^{\infty} \pi(E_{n-1}). \end{aligned} \quad (2.12)$$

Аналогично,

$$\begin{aligned} L_0 &= \sum_{n=0}^{\infty} \int A^*(B^*JA^*)^n C^* \cdot \varphi(\mathbf{v}) d\mathbf{v}, \\ L_1 &= \sum_{n=0}^{\infty} \int A^*(B^*JA^*)^n C^* \cdot \frac{\psi(\mathbf{v})}{\varphi(\mathbf{v})} \varphi(\mathbf{v}) d\mathbf{v}. \end{aligned} \quad (2.13)$$

Наконец, оказалось удобным для уменьшения статистического разброса перейти к модели, в которой акт поглощения и переход к энергии ниже пороговой не разыгрывается, а учитывается изменением массы. После каждого переноса изменяем массу по формуле

$$\tilde{\mu}_k = \frac{\Sigma_{\text{pac}}(E_{k-1})}{\Sigma(E_{k-1})} \mu_{k-1}, \quad \mu_0 = 1 \quad (k \geq 1),$$

и затем перед каждым рассеянием по формуле

$$\mu_k = [1 - \pi(E_{k-1})] \tilde{\mu}_k \quad (k \geq 1).$$

Чтобы рассчитать все числовые характеристики на одних и тех же траекториях, начальная энергия всегда разыгрывалась по индикатрисе источника. Поэтому при усреднении по индикатрисе деления  $U^{238}$  каждая траектория бралась с весом  $\psi(v_0)/\varphi(v_0)$ , где  $v_0(E_0)$  — разыгрываемая начальная скорость. На каждой траектории  $\xi$  вычислялись функционалы

$$\begin{aligned} \Phi_1[\xi] &= \tilde{l}_0 + \sum_{k=1}^{\infty} \tilde{\mu}_k \pi(E_{k-1}), \\ \Phi_2[\xi] &= \tilde{l}_1 + \frac{\psi(v_0)}{\varphi(v_0)} \sum_{k=1}^{\infty} \tilde{\mu}_k \pi(E_{k-1}), \\ \Phi_3[\xi] &= \sum_{k=1}^{\infty} \tilde{\nu}(E_k) \mu_k \frac{\Sigma_{\text{дел}}(E_k)}{\Sigma(E_k)}, \\ \Phi_4[\xi] &= \frac{\psi(v_0)}{\varphi(v_0)} \sum_{k=1}^{\infty} \tilde{\nu}(E_k) \mu_k \frac{\Sigma_{\text{дел}}(E_k)}{\Sigma(E_k)}. \end{aligned} \quad (2.14)$$

Их средние значения принимались в качестве оценок искомых величин  $l_0$ ,  $l_1$ ,  $L_0$ ,  $L_1$ .

По заданию Института атомной энергии им. И.В. Курчатова нами были проведены расчеты коэффициента размножения на быстрых нейтронах в среде из  $U^{238}$  (см. [1]). В расчетах была применена описанная выше методика. Для оценки точности кроме функционалов  $\Phi_i$  вычислялись также функционалы  $\overline{\Phi_i \Phi_j}$ , по которым строилась приближенная матрица вторых моментов  $\{\overline{\Phi_i \Phi_j}\}$ . Для повышения точности расчетов моделировались независимые группы связанных между собой траекторий (по  $v_0$  бралась квадратурная формула повышенной точности; ср., например, [7]). Описанная здесь схема моделирования была найдена в результате большого числа экспериментов. При этом возникали большие трудности, так как при изменении отдельных частей алгоритма задачу приходилось заново программировать (всего было рассмотрено более десяти вариантов программ).

Использование описанной методики позволило значительно сократить расчетное время. Так, например, при расчете  $\epsilon$  по индикатрисе источника  $f^{235}(E) = 0,484e^{-E} \operatorname{sh} \sqrt{2E}$ , индикатрисе деления  $f^{238} \sim e^{-\beta E} \operatorname{sh} \sqrt{2E}$ , ( $\beta = 1$ ), индикатрисе неупругого рассеяния на уране  $p(E', E) \sim Ee^{-E/T}$  ( $T = \sqrt{E/a}$ ,  $a = 12,38$ ) и переменном коэффициенте размножения  $\bar{\nu} = 2,57 + 0,12(E - 1,15)$  за один час работы машины “Стрела” было смоделировано  $\approx 7000$  траекторий и получено значение  $\epsilon = 1,277$  со средней квадратичной ошибкой 0,0009. Для сравнения укажем, что при прямом моделировании физического процесса для достижения той же точности потребовалось бы примерно в  $10^3$  раз больше машинного времени.

### 3. Расчет распространения гамма-излучения в воздухе

В безграничной однородной воздушной среде задан точечный источник, излучающий  $\gamma$ -кванты энергии  $E_0$  изотропно по всем направлениям. Требуется рассчитать интенсивность рассеянного излучения вокруг источника.

Движение фотона подчиняется законам релятивистской механики. Поэтому положение частицы должно характеризоваться не парой  $(\mathbf{x}, \mathbf{v})$ , а парой  $(\mathbf{x}, \mathbf{p})$ , где  $\mathbf{p}$  — импульс  $\gamma$ -кванта. На практике, однако, удобно задавать импульс не его составляющими, а энергией  $E$  и направлением  $\Omega$  движения фотона. Такой заменой часто удобно пользоваться и при моделировании движения нерелятивистских частиц, потому что характеристики среды, как правило, зависят только от  $|\mathbf{v}|$ , или, что то же самое, от энергии  $E = mv^2/2$ . Разумеется, при этом плотности в новых координатах отличаются от старых плотностей на якобиан перехода от старых координат к новым.

Добавим новую координату  $t$  — время, отсчитываемое от момента испускания фотона, расширив тем самым фазовое пространство  $\Omega$ . Очевидно, так определенное  $t$  является аддитивным функционалом от траектории. Движение частицы в расширенном фазовом пространстве  $\{\Omega, t\}$  также является марковским.

Далее, поскольку излучатель изотропный, поле излучения сферически симметрично. Его естественно характеризовать по пространству единственной координатой  $R$  — расстоянием от источника. Обозначим через

$$J(\mathbf{x}, E, \Omega, t) = J(R, E, \Omega, t), \quad (3.1)$$

где  $R = |\mathbf{x}|$ , плотность энергии, переносимой в момент  $t$  в точке  $\mathbf{x}$  по направлению  $\mathbf{\Omega}$ . Эта величина связана с плотностью переноса  $u(t, \mathbf{x}, \mathbf{p})$  и, как вытекает из формулы (1.2), с плотностью  $q(t, \mathbf{x}, \mathbf{p})$  соотношением

$$J(R, E, \mathbf{\Omega}, t) = Eu(t, \mathbf{x}, \mathbf{p}) = \frac{E}{c\Sigma(E)} q(t, \mathbf{x}, \mathbf{p}), \quad (3.2)$$

где  $c$  — скорость света.

При решении задачи помимо самой плотности  $J(E_0, R, E, \mathbf{\Omega}, t)$  вычислялись следующие величины:

$$J(E_0, R) = \int_0^\infty \oint \int_0^{E_0} J(E_0, R, E, \mathbf{\Omega}, t) dE d\mathbf{\Omega} dt$$

— полная интенсивность излучения в данной точке, т.е. полная интенсивность излучения, регистрируемого приемником на расстоянии  $R$  от постоянно действующего источника  $\gamma$ -квантов энергии  $E_0$ ;

$$J(E_0, R, \tau) = \oint \int_0^{E_0} J(E_0, R, E, \mathbf{\Omega}, t) dE d\mathbf{\Omega}$$

— полная интенсивность излучения в данной точке в момент  $\tau = t - R/c$ , т.е. суммарная интенсивность излучения на расстоянии  $R$  от  $\delta$  (источника  $\gamma$ -квантов) энергии  $E_0$  через время  $\tau$  после прохождения прямого излучения;

$$\begin{aligned} J(E_0, R, E) &= \int_0^\infty \oint J(E_0, E, \mathbf{\Omega}, t) d\mathbf{\Omega} dt; \\ J(E_0, R, \mathbf{\Omega}) &= \int_0^\infty \int_0^E J(E_0, R, E, \mathbf{\Omega}, t) dE dt; \\ J(E_0, R, E, \tau) &= \oint J(E_0, R, E, \mathbf{\Omega}, t) d\mathbf{\Omega}; \\ J(E_0, R, \mathbf{\Omega}, \tau) &= \int_0^{E_0} J(E_0, R, E, \mathbf{\Omega}, t) dE; \\ J(E_0, R, E, \mathbf{\Omega}) &= \int_0^\infty J(E_0, R, E, \mathbf{\Omega}, t) dt; \end{aligned}$$

$J(E_0, R, \tau, E, \mathbf{\Omega})$  — сама исходная величина  $J$ .

Простейшей характеристикой поля излучения является полная интенсивность рассеянного излучения — величина

$$J(E_0, R) - e^{-\Sigma(E_0)R},$$

т.е. суммарная интенсивность за вычетом интенсивности прямого излучения. С этой величиной связан фактор накопления

$$B_E(R) = J(E, R) e^{\Sigma(E)R}$$

— отношение полной интенсивности излучения к интенсивности прямого излучения.

Система интегральных уравнений (II) в данной задаче упрощается, поскольку “размножением” фотонов можно пренебречь<sup>4)</sup>:

$$\begin{cases} q = \tilde{A}(p + \varphi), \\ p = \tilde{B}q. \end{cases} \quad (3.3)$$

Исключив  $p$ , получим уравнение только для  $q$ :

$$q = ABq + A\varphi;$$

отсюда получим

$$q = [I + (AB) + \dots + (AB)^n + \dots] A\varphi = A[I + (AB) + \dots + (AB)^n + \dots] \varphi,$$

или, обозначив  $(BA) = L$ ,

$$q = A \sum_{n=0}^{\infty} L^n \varphi. \quad (3.4)$$

При розыгрыше процесса переноса в естественном направлении приходится вычислять не значения плотности  $q$  в избранных точках, а систему значений скалярных произведений  $(f, q)$  функции  $q$  с выбранными пробными функциями  $f$ . Так как нас интересует не плотность  $q$ , а выражаемая через нее интенсивность  $J$ , то мы выберем  $f = \frac{E}{c\Sigma} \psi$  так, чтобы  $(f, q) = (\psi, J)$ . Воспользуемся равенством

$$\left( f, A \sum_{n=0}^{\infty} L^n \varphi \right) = \left( A^* f, \sum_{n=0}^{\infty} L^n \varphi \right).$$

Ядро операторов  $A$  и  $A^*$  вырождено:

$$a(\mathbf{x}', E', \boldsymbol{\Omega}', t'; \mathbf{x}, E, \boldsymbol{\Omega}, t) = a(\mathbf{x}', \mathbf{x}) \cdot \delta_{E'}(E) \cdot \delta_{\boldsymbol{\Omega}'}(\boldsymbol{\Omega}) \cdot \delta_{t'}\left(t - \frac{|\mathbf{x}' - \mathbf{x}|}{c}\right),$$

где  $a(\mathbf{x}', \mathbf{x})$  — обобщенная плотность распределения, сосредоточенного на луче  $\mathbf{l} = \mathbf{x} + s\boldsymbol{\Omega}$ ,  $s \geq 0$ . Обыкновенный интегральный оператор  $M^*$  с ядром  $m(x, y)$  переводит обобщенную функцию  $\delta(x - x_0)$  в обыкновенную функцию  $m(x_0, y)$ . Для оператора  $A^*$  в силу его вырожденности это не так. Максимум, что удастся сделать, — это выбрать пробную функцию в виде

$$\psi = \chi_{\Delta} \cdot \delta_R,$$

где  $\delta_R$  —  $\delta$ -функция, сосредоточенная на сфере радиусом  $R$ , с единичной плотностью на поверхности<sup>5)</sup>, а  $\chi_{\Delta}$ , например, есть характеристическая функция многомерного интервала

$$\Delta = (E, E + \Delta E) \times (\tau, \tau + \Delta\tau) \times \Delta\boldsymbol{\Omega}.$$

<sup>4)</sup> В начале раздела была введена дополнительная координата  $t$ . Знак тильды поставлен над операторами  $A$  и  $B$ , чтобы подчеркнуть, что операторы  $\tilde{A}$  и  $\tilde{B}$  описывают перенос в расширенном фазовом пространстве  $\{\mathbf{x}, \mathbf{p}, t\}$ . Случайное блуждание по присоединенной координате является вырожденным. Соответственно ядро оператора  $\tilde{A}$  отличается от ядра оператора  $A$  на множитель  $\delta(t - |\mathbf{x}' - \mathbf{x}|/c - t')$ , а ядро оператора  $\tilde{B}$  отличается от ядра  $B$  на  $\delta(t - t')$ . Поскольку это не приведет к недоразумениям, знак тильды в дальнейшем будем опускать.

<sup>5)</sup> При таком выборе  $(\psi_{\Delta}, J) = 4\pi R^2 \int_E^{E+\Delta E} \int_{\tau}^{\tau+\Delta\tau} \int_{\Delta\boldsymbol{\Omega}} J(R, E, \boldsymbol{\Omega}, t) d\boldsymbol{\Omega} dt dE = 4\pi R^2 J_{\Delta}(R)$ .

Интеграл  $A^* \psi$  от обобщенной функции  $\psi$  удается вычислить аналитически через значения некоторой функции  $q$  в нескольких точках. Проведем выкладки:

$$\begin{aligned} A^* f &= \int_t \int_{\mathbf{x}} \int_E \oint_{\Omega} \left[ c \Sigma e^{-\int_0^s \Sigma d\sigma} \cdot \frac{\delta(\omega - \Omega')}{s^2} \cdot \delta(\Omega - \Omega') \delta\left(t - \frac{s}{c} - t'\right) \cdot \delta(E - E') \frac{E}{c\Sigma} \right] \times \\ &\quad \times [\bar{\delta}(r - R) \cdot \chi_{\Delta}(\Omega, E, \tau)] d\Omega dE d\mathbf{x} dt = \\ &= \int_t \int_{\mathbf{x}} e^{-\int_0^s \Sigma d\sigma} \cdot \frac{\delta(\omega - \Omega')}{s^2} \delta\left(t - \frac{s}{c} - t'\right) E' \delta(r - R) \chi_{\Delta}(\Omega', E', \tau) d\mathbf{x} dt, \end{aligned}$$

где  $s = |\mathbf{x} - \mathbf{x}'|$ ,  $\omega = (\mathbf{x} - \mathbf{x}')/s$ .

В записи интеграла использованы две системы полярных координат:  $r, \Theta$  с центром в источнике и  $(s, \omega)$  с центром в точке  $\mathbf{x}$ . Перейдем от декартовой системы координат к смешанной системе  $(r, \omega)$ . Якобиан перехода равен  $s^2/|\cos \nu|$ , где  $\cos \nu = (\Theta, \omega)$  (геометрически  $\nu$  — угол между направлениями луча  $\mathbf{x} + s\omega$  и радиусом  $\mathbf{r}$  сферы в точке пересечения луча со сферой). Соответствие между декартовой системой и новой не однозначное: в зависимости от положения точки  $\mathbf{x}'$  значению  $(r, \omega)$  могут соответствовать две точки, одна точка или не будет соответствовать ни одна точка (пересечения сферы радиуса  $r$  с лучом  $\mathbf{x} + s\omega$ ). Отсюда

$$\begin{aligned} A^* f &= \int_t \oint_{\omega} \int_{\mathbf{r}} e^{-\int_0^{s(r, \omega)} \Sigma d\sigma} \frac{E'}{|\cos \nu|} \delta(\omega - \Omega') \times \\ &\quad \times \delta\left(t - \frac{s(r, \omega)}{c} - t'\right) \delta(r - R) \chi_{\Delta}(\Omega', E', \tau) dr d\omega dt = \\ &= \sum_{s(R, \Omega')} \int_t e^{-\int_0^{s(R, \Omega')} \Sigma d\sigma} \frac{E'}{|\cos \nu|} \delta\left(t - \frac{s(R, \Omega')}{c} - t'\right) \chi_{\Delta}(E', \Omega', \tau) dt = \\ &= \sum_{s(R, \Omega')} \frac{E'}{|\cos \nu(R, \Omega')|} e^{-\int_0^{s(R, \Omega')} \Sigma d\sigma} \chi_{\Delta}\left(E', \Omega', t' + \frac{s(R, \Omega') - R}{c}\right), \quad (3.5) \end{aligned}$$

где сумма берется по всем точкам  $\mathbf{x}$  пространства с координатами  $(R, \Omega')$ . Назовем выражение для  $A^* f$  функцией  $G$ . Тогда

$$A^* f = G_{\Delta}(\mathbf{x}', E', \Omega', t'; R) = \sum_{s(R, \Omega')} g_{\Delta}(x', E', \Omega', t'; R) \quad (3.6)$$

(вид функции  $g$  очевиден).

В процессе переноса  $\gamma$ -квантов возможны два типа взаимодействия с веществом: рассеяние и поглощение. При стохастическом моделировании поглощение учитывалось изменением массы частицы. Вероятность рассеяния тогда выделяется из оператора  $\mathbf{B}$  в качестве фактора  $z(x, y)$ . Таким образом, моделировалось движение марковской частицы с массой  $\mu_k$ , изменяющейся по закону  $\mu_k = (\Sigma_{\text{рас}}/\Sigma)\mu_{k-1}$  ( $\mu_0 = 1$ ), рассеивающейся, как и фотон, по индикатрисе Клейна-Нишины и с прежним полным сечением взаимодействия с веществом.

На траекториях частицы вместо функционалов

$$\sum_{k=0}^{\nu-1} G_{\Delta}(\mathbf{x}_k, E_k, \boldsymbol{\Omega}_k, t_k; R),$$

средние значения которых по траекториям физического процесса равны искомым величинам

$$4\pi R^2 J_{\Delta}(R) = (f, q) = ([I - A^*B^*]^{-1}A^*f, \varphi),$$

вычислялась система функционалов

$$\Phi_{\Delta, R}[\xi] = \sum_{k=0}^{\infty} G_{\Delta}(\mathbf{x}_k, E_k, \boldsymbol{\Omega}_k, t_k; R)\mu_k. \quad (3.7)$$

Траектории исходного процесса с вероятностью единица кончаются на некотором случайном  $\nu$ -м положении, когда происходит захват фотона. Траектории моделируемого процесса бесконечны, поэтому их обрывали, как только  $\mu_k$  станет меньше выбранного  $\varepsilon$ .

Средние величины  $\overline{\Phi_{\Delta, R}}$  приближенно оценивают искомую интенсивность  $I_{\Delta}(R)$ . Остальные величины получаются суммированием  $\overline{\Phi_{\Delta, R}}$  по соответствующим переменным.

Для решения задачи были составлены программы для машин “Стрела” и М-20. Применялась блочная система программирования. В процессе работы производились неоднократные изменения и совершенствование методики и программы. В связи с этим изменялись также и отдельные программные блоки, а общая программа заново компоновалась из частей с помощью компоновочной программы.

Исходные данные для решения задачи принимались следующие: плотность воздуха  $1,29 \cdot 10^{-3}$  г/см<sup>3</sup>, состав воздуха — 78% азота, 21% кислорода и 1% аргона; значения полного коэффициента поглощения, сечений фотоэффекта, образования пар и комптоновских рассеяний были пересчитаны по данным, приведенным в работе [8].

В программе перенос излучения одновременно регистрировался на пяти расстояниях  $R_k$ , что позволило примерно в три раза сократить общий объем вычислений.

В результате расчета выдавались трехмерные гистограммы величин  $I_{\Delta}(R_k)$  и их свертки. По заданию Института химической физики АН СССР было сосчитано большое число вариантов при начальных энергиях источника  $E_0 = 0,5; 1,0; 2,0; 4,0; 6,0$  МэВ.

При проведении расчета непосредственно считались среднеквадратичные ошибки полной энергии, проносимой излучением через сферу радиуса  $R_k$ , ошибки энергетических распределений интенсивности и ошибки временного распределения интенсивности. Ошибки остальных величин оценивались по последовательным выдачам данных с машины (по  $N = 500, 1000, 1500$  траекторий). Кроме того, проводились сравнительные расчеты одних и тех же вариантов с разными последовательностями псевдослучайных чисел. Анализ ошибок показал,

что выбранная методика расчетов даже по сравнительно малому числу траекторий ( $N = 2000$ ; время счета варианта при такой статистике — 15–20 ч работы машины “Стрела”) обеспечивает вычисление полной интенсивности рассеянного излучения  $I_{\Delta}(R)$  при  $\Sigma R < 5$  с точностью не хуже 10%. Точность временных распределений интенсивности в диапазоне времени до 1 мкс составляет 10–20%, после 1 мкс 20–40% с ухудшением точности на больших расстояниях. Угловые энергетические спектры и временные спектры считаются при указанной статистике с точностью 15–30%.

На рис. 1 и 2 сопоставляются результаты, полученные методом Монте-Карло, с результатами для постоянно действующего источника, сосчитанные методом моментов в работе [9] (в этой работе рассчитывалось распространение  $\gamma$ -квантов в воде; в рассматриваемом диапазоне энергий данные для воздуха и воды близки при сравнении их в безразмерном масштабе расстояний  $\Sigma(E_0)R$ ).

На рис. 1 виден плавный характер изменения вычисленных значений  $B_{E_0}(R)$

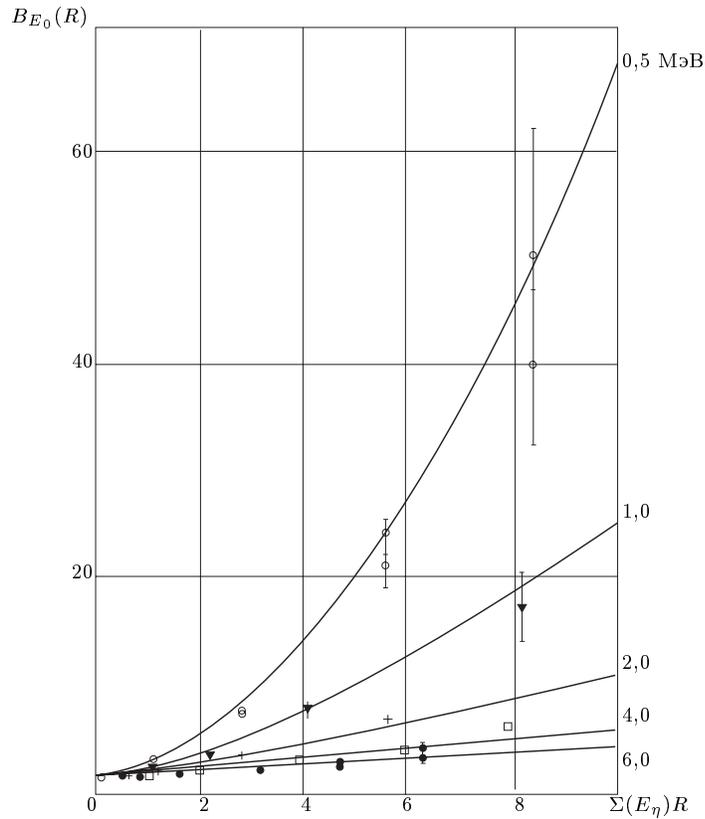


Рис. 1. Зависимость фактора накопления  $B_{E_0}(R)$  от расстояния  $\Sigma(E_0)R$  для различных начальных энергий  $E_0$ : сплошные кривые сосчитаны методом моментов, значками обозначены результаты расчетов методом Монте-Карло (для  $E = 0,5$  МэВ нанесены данные двух серий “наблюдений”)

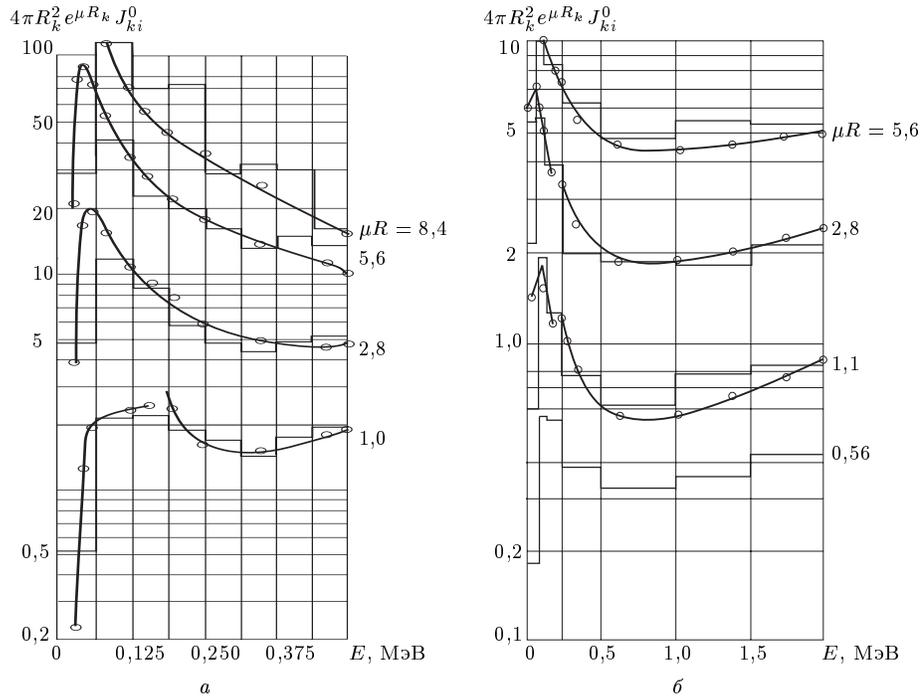


Рис. 2. Гистограммы дифференциальных энергетических спектров интенсивности  $J(E_0, R, E)$  для различных расстояний от источника: гладкие кривые — спектры, сосчитанные по методу моментов; а —  $E_0 = 0,5$  МэВ; б —  $E_0 = 2,0$  МэВ

по расстоянию и энергии внутри каждой серии. На графиках рис. 2 наблюдается вполне удовлетворительное совпадение результатов.

#### 4. Расчет процессов рассеяния и захвата нейтронов в бороводородном сцинтилляторе

Сцинтиллятор представляет собой сферу или цилиндр, заполненные однородной смесью водорода, углерода и бора<sup>6)</sup>. По диаметру сферы или по оси цилиндра падает поток нейтронов энергии  $E_0$ . Требуется определить следующие характеристики сцинтиллятора:

эффективность  $u$  — вероятность поглощения нейтрона бором при условии предварительного взаимодействия с ядром водорода ( $u$  является основной характеристикой прибора);

$\{\tau\}$  — распределение времен от момента первого рассеяния нейтрона на водороде до момента его поглощения ядром бора;

$\{\mathcal{E}\}$  — распределение энергий нейтрона в момент его поглощения ядром бора при условии предварительного взаимодействия нейтрона с ядром водорода;

<sup>6)</sup>В состав вещества реального сцинтиллятора входит также кислород, однако в задаче его присутствие отдельно не учитывалось.

$\{\Delta_E\}$  — распределение суммарной потери энергии нейтрона при его рассеяниях на ядрах водорода (для нейтронов, поглощенных бором).

Для простоты предполагалось, что рассеяние происходит на ядрах водорода и углерода и притом изотропно в лабораторной системе координат, а ядра бора только поглощают.

Расширим фазовое пространство, добавив координату времени  $t$ , а также номер элемента, с ядром которого происходит взаимодействие.

Система интегральных уравнений переноса имеет следующий вид:

$$\begin{aligned} q &= A(p_1 + p_2 + p_3 + \varphi), \\ p_1 &= B_1 \pi_1 q, \\ p_2 &= B_2 \pi_2 q, \\ p_3 &= 0 \cdot \pi_3 q, \end{aligned} \quad (4.1)$$

где  $p_1$  — плотность числа нейтронов, рассеянных на ядрах углерода;  $p_2$  — плотность числа нейтронов, рассеянных на ядрах водорода;  $p_3$  — равная нулю плотность числа нейтронов, рассеянных на ядрах бора;  $\pi_1(E)$  — вероятность взаимодействия (рассеяния) с углеродом;  $\pi_2(E)$  — вероятность взаимодействия (рассеяния) с водородом;  $\pi_3(E)$  — вероятность взаимодействия (поглощения) с бором.

В “реальном” физическом процессе при взаимодействии нейтрона с веществом каждый раз происходит “выбор” с вероятностями  $\pi_1, \pi_2, \pi_3$  ядра, с которым произошло взаимодействие.

Положим  $B = B_1 \pi_1 + B_2 \pi_2 + 0 \pi_3$  и формально выпишем

$$q = [I + AB + \dots + (AB)^n + \dots] A \varphi.$$

Интересующая нас плотность поглощения бором равна

$$\pi_3 q = \pi_3 A [I + BA + \dots + (BA)^n + \dots] \varphi. \quad (4.2)$$

Плотность распределения вероятностей траектории движения частицы в расширенном фазовом пространстве получается следующим образом. Член  $\pi_3 A (BA)^n$  описывает долю частиц, испытавших  $n$  рассеяний и затем поглощенных бором. Ее можно представить в виде

$$\pi_3 A (BA)^n = \sum_{\gamma_1, \dots, \gamma_n=1}^2 \pi_3 A_{n+1} B_{\gamma_n} \pi_{\gamma_n} A_n \dots B_{\gamma_1} \pi_{\gamma_1} A_1 \varphi, \quad (4.3)$$

где каждое слагаемое описывает вклад в  $q$  от траекторий, по которым нейтрон последовательно взаимодействовал с элементами  $\gamma_1, \gamma_2, \dots, \gamma_n$ . Соответствующая плотность аналогична (4.3), только вместо операторов выписываются соответствующие ядра. Искомые характеристики записываются как средние значения функционалов от траектории частицы. Обозначим:  $\gamma = (\gamma_1, \dots, \gamma_\nu)$  — случайная последовательность номеров элементов, с ядрами которых происходит взаимодействие,  $\gamma_s = 0, 1, 2, 3$  (0 — индекс вылета из сцинтиллятора);

$$i(\gamma) = \begin{cases} i & \text{при } \gamma_i = 2, \gamma_1 = \gamma_2 = \dots = \gamma_{i-1} = 1, \\ 0 & \text{при } \gamma_1 = \gamma_2 = \dots = \gamma_{\nu-1} = 1, \quad \gamma_\nu = 0 \text{ или } \gamma_\nu = 3. \end{cases}$$

Тогда

$$\begin{aligned}
 u &= \begin{cases} 1, & \gamma_\nu = 3, \\ 0, & \gamma_\nu = 0, \end{cases} & (4.4) \\
 \tau &= t_\nu - t_{i(\gamma)}, \\
 \mathcal{E} &= \begin{cases} E_{\nu-1}, & i(\gamma) < \nu, \quad \gamma_\nu = 3, \\ 0 & \text{во всех остальных случаях,} \end{cases} \\
 \Delta_E &= \begin{cases} \sum_{j: \gamma_j=2} (E_{j-1} - E_j), & \gamma_\nu = 3, \\ 0 & \text{во всех остальных случаях.} \end{cases}
 \end{aligned}$$

Соответствующие кратные интегралы вычислялись методом Монте-Карло с помощью следующей стохастической модели. При моделировании переноса по пространству (оператор  $A$ ) была применена схема условного вылета, в которой разыгрывалось только положение точки очередного взаимодействия частицы внутри сцинтиллятора и вычислялась соответствующая потеря массы. Кроме того, поглощение также учитывалось изменением массы. На каждом столкновении по-прежнему случайно выбирался тип рассеяния. Изменение массы задавалось парой рекуррентных соотношений

$$\begin{aligned}
 \tilde{\mu}_k &= (1 - W_{k-1})\mu_{k-1} \quad (\mu_0 = 1), \\
 \mu_k &= (1 - \pi_3)\tilde{\mu}_k,
 \end{aligned} \tag{4.5}$$

где  $W_k$  — вероятность вылета после  $k$ -го взаимодействия.

При переходе от физического процесса переноса к новой модели блуждания исходные функционалы, стоящие под знаком интеграла, преобразуются соответственно в следующие:

$$\begin{aligned}
 \Phi_1[\xi] &= \sum_{n=i(\gamma)+1}^{\infty} \pi_3(E_{n-1})\tilde{\mu}_n, \\
 \Phi_2[\xi] &= \sum_{n=i(\gamma)+1}^{\infty} (t_n - t_{i(\gamma)})\pi_3(E_{n-1})\tilde{\mu}_n, \\
 \Phi_3[\xi] &= \sum_{n=i(\gamma)+1}^{\infty} E_{n-1}\pi_3(E_{n-1})\tilde{\mu}_n, \\
 \Phi_4[\xi] &= \sum_{n=i(\gamma)+1}^{\infty} \left[ \sum_{j: \gamma_j=2} (E_{j-1} - E_j) \right] \pi_3(E_{n-1})\tilde{\mu}_n.
 \end{aligned} \tag{4.6}$$

Гистограммы распределений величин  $\tau$ ,  $\mathcal{E}$ ,  $\Delta_E$  получаются усреднением функционалов  $\Phi_l^{(k)} = \Phi_l \chi_l^{(k)}$ , где  $\chi_l^{(k)}$  — характеристическая функция  $k$ -го интервала группировки для соответствующей величины. Кроме того, аналогично вычислялись многомерные гистограммы указанных величин. Разумеется, как и в разделе 3, траектории прослеживались до некоторого момента  $\nu$ , где  $\mu_\nu < \varepsilon$ .

Задача решалась на машине М-20. Окончательным расчетам предшествовал большой числовой эксперимент. Были изучены последовательности типов взаимодействий — “слова”  $\gamma$ , отвечающие классам траекторий, дающие основной вклад в эффективность  $u$  (так сказать, высоковероятные цепочки). Можно было ожидать, что при вычислении искомых интегралов отдельно по классам траекторий общий объем расчетов уменьшится. Эксперимент показал, что достаточно разделить траектории на два класса: с первым столкновением на водороде и с первым столкновением на углероде. Поэтому в окончательной расчетной схеме тип первого столкновения не разыгрывался, а траектория ветвилась на две: первой, соответствующей рассеянию на углероде, приписывалась масса  $\mu_1 = [1 - \pi_3(E_0)] \pi_1(E_0) \tilde{\mu}_1$ ; второй —  $\mu_1 = [1 - \pi_3(E_0)] \pi_2(E_0) \tilde{\mu}_1$ . Обе ветви моделировались на одной и той же последовательности случайных чисел, и их общий вклад сразу же суммировался.

Было рассчитано три состава сцинтиллятора соответственно из смеси:

- 1) ксилола ( $C_8H_{10}$ ) и триметилбората  $[B(OCH_3)_3]$  (необогатенный бор,  $B^{10}$  18%);
- 2) фенилциклогексана ( $C_{10}H_{16}$ ) и триизобутилбората  $[B(OC_4H_9)_3]$  (необогатенный бор);
- 3) ксилола ( $C_8H_{10}$ ) и триметилбората  $[B(OCH_3)_3]$  (обогатенный до 80% по  $B^{10}$  бор).

Подробные данные и результаты расчетов приведены в работах [3, 10].

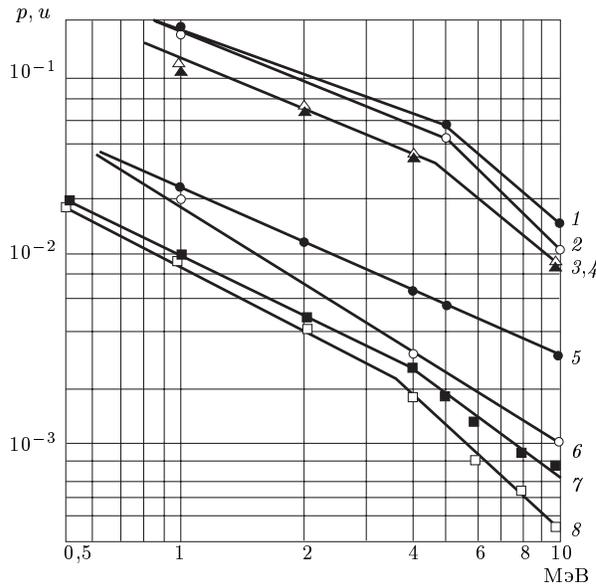


Рис. 3. Зависимость эффективности  $u$  и полной вероятности захвата  $p$  от начальной энергии: 1, 3, 5, 7 — величина  $p$ ; 2, 4, 6, 8 — величина  $u$

Было изучено 252 типа форм и размеров сцинтилляторов (сферы диаметром 80, 50, 40 мм и цилиндры диаметром 80, 50, 40 мм с такой же высотой) при энергиях пучка нейтронов  $E_0 = 0,5; 1,0; 2,0; 4,0; 6,0; 8,0; 10,0$  МэВ.

Всего было рассчитано около 50 производственных вариантов. Время расчета каждого в среднем составляло 20–40 мин. Достигнутая при этом относительная точность в ведущей характеристике — эффективности — составляла 10–15 %.

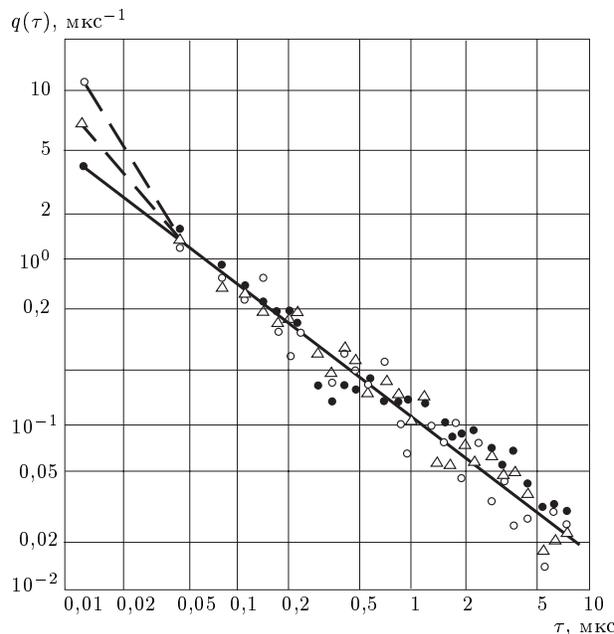


Рис. 4. Плотность  $q(\tau)$  распределения величины  $\tau$  (сравнение гистограмм с эмпирическим законом;  $D = 40$  мм, 1-й состав):  $\bullet$  —  $E_0 = 1$  МэВ,  $\Delta$  —  $E_0 = 4$  МэВ,  $\circ$  —  $E_0 = 10$  МэВ

Вероятность поглощения нейтронов в сцинтиляторе чрезвычайно мала. Можно с уверенностью сказать, что в этих условиях использование прямого моделирования неприемлемо. Еще одним фактором значительного уменьшения объема вычислений явилось использование зависимых испытаний [11–24]. Оказалось, что сосчитанные величины изменяются достаточно гладко (в частности, это обстоятельство позволило сократить общее число рассчитываемых вариантов). Были получены гладкие кривые зависимостей эффективности сцинтилятора от начальной энергии и размеров<sup>7)</sup>. На рис. 3 и 4 приведены наиболее интересные зависимости.

На рис. 3 показана зависимость эффективности  $u$  и полной вероятности захвата  $p$  от начальной энергии. На графике отмечены отдельные значения этих величин, а также кривые эмпирических зависимостей для следующих вариантов:

Вариант	$D$ , мм	Состав
1,2	80	2
3,4	80	3
5,6	40	2
7,8	40	1

<sup>7)</sup> Любопытно отметить, что в контрольных расчетах по независимым сериям испытаний случайные флуктуации почти полностью маскировали характер функциональных зависимостей.

На рис. 4 показана плотность  $q(\tau)$  распределения величины  $\tau$  (гистограммы сравниваются с эмпирическим законом). Нанесены значения плотности захватов, взятые из вычисленных гистограмм для трех вариантов. Из графика видно, что время от момента первого рассеяния на водороде до момента захвата бором практически не зависит от начальной энергии (вплоть до  $\tau = 0,05$  мкс). Было замечено, что зависимость эффективности  $u$  от  $E_0$  и  $V$  (объема сцинтиллятора) в основной области начальных энергий хорошо описывается эмпирическим законом

$$u = C \frac{V}{E_0},$$

где  $C$  — константа, зависящая от состава сцинтиллятора.

На основе результатов расчетов был спроектирован и построен нейтронный спектрометр с бороводородным сцинтиллятором. В контрольных измерениях спектров, проведенных на потоках нейтронов деления, получены результаты, близкие к известным.

#### СПИСОК ЛИТЕРАТУРЫ

1. Гельфанд И.М. и др. // Труды Второй международной конференции по мирному использованию атомной энергии. Женева, 1958. Доклады советских ученых. Т. 2. — М.: Атомиздат, 1959. — С. 628.
2. Лейпунский О.И. и др. // Атомная энергия. 1961. Т. 10, вып. 5. С. 493.
3. Аваев В.Н. и др. // Приборы и техника эксперимента. 1963. № 4. С. 39.
4. Symposium on Monte Carlo Methods / Ed. H. A. Meyer. — N.Y.: Wiley, 1956. — P. 89, 146.
5. Глестон С., Эдлунд М. Основы теории ядерных реакторов. — М.: ИЛ, 1954.
6. Дуб Дж.Л. Вероятностные процессы. — М.: ИЛ, 1956.
7. Гельфанд И.М. и др. // Изв. вузов. Сер. матем. 1958. № 5. С. 32
8. Зигбан К. Бета- и гамма-спектроскопия. — М.: ИЛ, 1959. — Приложение 1.
9. Goldstein H., Wilkins L. // Rept. U. S. Atomic Energy Comm. 1955. № 40. P. 3075
10. Аваев В.Н. и др. // Вопросы физики защиты реакторов. — М.: Госатомиздат, 1963. — С. 289.
11. Фролов А.С., Ченцов Н.Н. // Труды VI Всесоюзного совещания по теории вероятностей и математической статистике. — Вильнюс: Гос. изд. полит. и научн. лит. ЛитССР, 1962. — С. 425.
12. Бусленко Н.П. и др. Метод статистических испытаний (Метод Монте-Карло). — М.: Физматгиз, 1962.
13. Бусленко Н.П., Шрейдер Ю.А. Метод статистических испытаний (Монте-Карло) и его реализация на цифровых вычислительных машинах. — М.: Физматгиз, 1961.
14. Крамер Г. Математические методы статистики. — М.: ИЛ, 1948.
15. Лейпунский О.И. и др. Распространение гамма-квантов в веществе. — М.: Физматгиз, 1960.
16. Лозв М. Теория вероятностей. — М.: ИЛ, 1962.
17. Файнштейн А. Основы теории информации. — М.: ИЛ, 1960.
18. Владимиров В.С. // Теория вероятн. и ее примен. 1956. Т. 1, вып. 1. С. 113.
19. Гельфанд И.М., Ченцов Н.Н. // Ж. эксперим. и теор. физ. 1956. Т. 31, № 6. С. 1106.
20. Кертис Д. // Успехи матем. наук. 1957. Т. 12, вып. 1. С. 149.
21. Коробов Н.М. // Докл. АН СССР. 1957. Т. 115, № 6. С. 1062.
22. Berger M.I. // Res. Nat Bur. Standards. 1955. V. 55, № 6. P. 343.
23. Hammersley I.M., Morton K.W. // Proc. Cambridge Philos. Soc. 1956. V. 52. P. 449.
24. Фролов А.С., Ченцов Н.Н. // Ж. вычисл. матем. и матем. физ. 1962. Т. 4. С. 714.

## ПСЕВДОСЛУЧАЙНЫЕ ЧИСЛА ДЛЯ МОДЕЛИРОВАНИЯ МАРКОВСКИХ ЦЕПЕЙ<sup>\*)</sup>

Как известно (см. [1, 2]), вычисление конфигурационных средних в задачах статистической физики может быть сведено к вычислению методом Монте-Карло математических ожиданий соответствующих величин по стационарному распределению соответственно подобранной цепи Маркова. В настоящей работе изучены условия, которым должна удовлетворять псевдослучайная последовательность чисел  $\xi_k$ ,  $0 \leq \xi_k \leq 1$ ,  $k = 1, 2, \dots$ , чтобы, какова бы ни была эргодическая цепь Маркова, при моделировании этой цепи с помощью  $\{\xi_k\}$  по общепринятым формулам “розыгрыша” исходов получалась бы “типичная”, т.е. эргодичная траектория, временное среднее вдоль которой асимптотически совпадало бы со средним по стационарному распределению цепи. Выяснено, что класс таких последовательностей совпадает с классом вполне равномерно распределенных последовательностей, введенных в [3]. Показано также, что обычные (см. [4, 5]) тесты и проверки только на равномерную распределенность недостаточны уже в простейших случаях.

1. Пусть  $\{\omega_k\}_{k=1}^{\infty}$  — последовательность независимых случайных точек  $\omega \in \Omega$  с распределением  $\mu\{d\omega\}$  каждая. Усиленный закон больших чисел утверждает (см. [6]), что для любой абсолютно интегрируемой по Лебегу функции  $f(\omega)$

$$\mathbf{P}\left\{\frac{1}{N} \sum_{k=1}^N f(\omega_k) \xrightarrow{N \rightarrow \infty} \int_{\Omega} f(\omega) \mu\{d\omega\}\right\} = 1. \quad (1)$$

Скорость сходимости к интегралу может быть оценена с помощью центральной предельной теоремы. Свойство (1) используется в методе Монте-Карло для приближенного вычисления интегралов. Если вероятностное пространство  $(\Omega, S, \mu)$  конструктивно, т.е. существует измеримое отображение  $\omega = \varphi(x)$  отрезка  $E = [0 \leq x \leq 1]$  в  $\Omega$  такое, что для любого  $A \in S$

$$\mu\{A\} = \int_B dx, \quad B = \{x : \varphi(x) \in A\}, \quad (2)$$

то

$$\int_{\Omega} f(\omega) \mu\{d\omega\} = \int_0^1 f(\varphi(x)) dx$$

и для вычислений можно воспользоваться последовательностью  $\{\xi_k\}$  независимых, равномерно распределенных на  $E$  случайных чисел.

В реальных расчетах на ЭВМ вместо случайной последовательности удобно брать псевдослучайную последовательность чисел, генерированную каким-либо арифметическим способом. Поэтому нужны критерии “пригодности” таких последовательностей, показывающие, может ли данная последовательность  $\{\xi_k\}$

<sup>\*)</sup> ЖВМ и МФ. 1967. Т. 7, № 3. С. 632–643.

считаться “типичным” представителем “основной массы” последовательностей в смысле (1), т.е. обеспечивать сходимость

$$\frac{1}{N} \sum_{k=1}^N f(\xi_k) \xrightarrow{N \rightarrow \infty} \int_0^1 f(x) dx. \quad (3)$$

К сожалению, не существует универсальной последовательности, обеспечивающей сходимость (3) для любой интегрируемой по Лебегу функции  $f(x)$ . Чтобы убедиться в этом, возьмем для примера функцию  $f(x) = 0$  при  $x = \xi_k$ ,  $k = 1, \dots, N, \dots$ , и  $f(x) = 1$  при  $x \notin \{\xi_k\}$ . Для нее интеграл равен 1, а все интегральные суммы с узлами в  $\xi_k$  равны 0. Поэтому приходится предъявлять к последовательности значительно более слабые требования. Например, при вычислении однократных интегралов обычно требуют равномерной распределенности последовательности  $\{\xi_k\}$  по Вейлю, которая эквивалентна сходимости (3) для любой функции  $f(x)$ , интегрируемой по Риману (см. [7, 8]). В качестве критерия пригодности здесь может быть взят критерий согласия Колмогорова или критерий  $\omega^2$ . При вычислении кратных интегралов

$$\int_0^1 \int_0^1 f(x, y) dx dy, \quad \int_0^1 \int_0^1 \int_0^1 f(x, y, z) dx dy dz$$

и т.д. координаты очередной счетной точки предпочитают формировать не из одного и того же числа, а каждую — из отдельного числа последовательности. Это объясняется тем, что на ЭВМ каждое число записывается своими первыми разрядами, всего не более 40 двоичных разрядов, и при отображении отрезка  $E$  в куб дискретность выбираемых точек оказывается существенной<sup>1)</sup>. Поэтому проверяют равномерную распределенность в единичном квадрате или в единичном кубе (соответствующей размерности) последовательности точек  $\eta_k$  с координатами

$$(\xi_1, \dots, \xi_s), (\xi_2, \dots, \xi_{s+1}), \dots, (\xi_k, \dots, \xi_{k+s-1}), \dots \quad (4)$$

или последовательности точек с координатами

$$(\xi_1, \dots, \xi_s), (\xi_{s+1}, \dots, \xi_{2s}), \dots, (\xi_{ks+1}, \dots, \xi_{(k+1)s}), \dots \quad (5)$$

в зависимости от того, каким способом из последовательности  $\{\xi_j\}$  мы получаем последовательность  $s$ -мерных векторов  $\eta_k$ , ср. [4].

Последовательность точек  $\eta_k = (\eta_{k,1}, \dots, \eta_{k,s})$  называется *равномерно распределенной*, если для любой интегрируемой по Риману (или, что дает эквивалентное определение, для любой непрерывной) функции  $f$  в единичном кубе  $E_s$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N f(\eta_{k,1}, \dots, \eta_{k,s}) = \int_0^1 \dots \int_0^1 f(y_1, \dots, y_s) dy_1 \dots dy_s. \quad (6)$$

<sup>1)</sup> Например, при простейшем отображении  $y = \varphi(x)$  по формуле  $y_i = \varphi_i(x) = 0, \alpha^{(i)} \alpha^{(i+s)} \alpha^{(i+2s)} \dots$ , где  $0, \alpha^{(1)} \alpha^{(2)} \alpha^{(3)} \dots$  — двоичное (или  $m$ -ичное) разложение числа  $x$ .

Свяжем с последовательностью  $\eta_k$  последовательность нормированных дискретных мер  $\mu_N$ ,  $\mu_N\{\eta_k\} = N^{-1}$  при  $k \leq N$ ,  $\mu_N\{E_s\} = 1$ . Легко видеть, что последовательность  $\{\eta_k\}$  будет равномерно распределена в  $E_s$  тогда и только тогда, когда  $\mu_N$  слабо сходится к равномерному распределению на  $E_s$ . О быстроте этой сходимости можно судить по величине уклонения эмпирической функции распределения

$$F_N^*(z_1, \dots, z_s) = \mu_N\{(y_1, \dots, y_s) : 0 \leq y_i \leq z_i, i = 1, \dots, s\}$$

от функции  $z_1 \cdot \dots \cdot z_s$ , задающей равномерное распределение. В качестве критериев здесь обычно применяют  $s$ -мерные аналоги критериев Колмогорова или Мизеса–Смирнова. Как известно (см. [9]), для гладкой функции  $f$  знание меры уклонения эмпирического распределения от теоретического позволяет дать детерминированную оценку уклонения среднего арифметического от искомого интеграла.

2. Пусть  $(p_{ij})_{i,j=1}^n$  — стохастическая матрица,

$$p_{ij} \geq 0, \quad \sum_{j=1}^n p_{ij} = 1, \quad i = 1, \dots, n.$$

Как известно (см. [10]), у всякой стохастической матрицы существует хотя бы один неотрицательный нормированный стационарный вектор  $(q_j)_{j=1}^n$ :

$$\sum_{i=1}^n q_i p_{ij} = q_j, \quad j = 1, \dots, n. \quad (7)$$

Обозначим через  $\omega^{(i)}$ ,  $i = 1, \dots, n$ , возможные исходы и через  $\nu_k$ ,  $k = 1, \dots, N, \dots$ , — последовательность случайных исходов, связанных в цепь Маркова с матрицей  $(p_{ij})$ :

$$\mathbf{P}\{\nu_{k+1} = \omega^{(i)} \mid \nu_k = \omega^{(i)}\} = p_{ij},$$

и с начальным распределением  $(r_i)$ :

$$\mathbf{P}\{\nu_0 = \omega^{(i)}\} = r_i.$$

Обычно моделирование такой цепи проводят следующим стандартным способом: задаются последовательностью независимых случайных чисел  $\xi_k$ ; по величине  $\xi_0$  и начальному распределению  $\mathbf{r} = (r_1, \dots, r_n)$  разыгрывают начальное положение  $\nu_0$ ; далее, зная  $(k-1)$ -е положение  $\nu_{k-1} = \omega^{(i)}$ , разыгрывают  $\nu_k$ , для чего сравнивают  $\xi_k$  с последовательными суммами  $\sum_{j=1}^m p_{ij}$ ,  $m = 0, 1, \dots, n$ ; если

$$\sum_{j=1}^{m-1} p_{ij} \leq \xi_k < \sum_{j=1}^m p_{ij}, \quad (8)$$

то полагают  $\nu_k = \omega^{(m)}$ ; начальное положение определяют по  $\xi_0$  аналогично. Нетрудно проверить по индукции простым подсчетом совместных вероятностей, что такой способ розыгрыша действительно дает марковскую последовательность с матрицей перехода  $(p_{ij})$  и начальным распределением  $(r_i)$ .

Пусть нам требуется вычислить среднее значение  $\bar{\Phi}$  функционала  $\Phi$  от марковской траектории, зависящего только от первых  $s$  исходов:  $\Phi = \Phi(\nu_0, \dots, \nu_{s-1})$ . Формулы (8) задают измеримое по Жордану отображение  $\varphi$  единичного  $s$ -мерного куба  $E_s$  в множество траекторий  $\{\nu_0, \dots, \nu_{s-1}\}$  и сводят задачу вычисления среднего  $\bar{\Phi}$  к вычислению  $s$ -кратного интеграла от интегрируемой по Риману функции  $\Phi(\varphi(\cdot))$ . Таким образом, справедлива

ЛЕММА 1. Пусть  $\{\eta_t\}$ ,  $\eta_t = (\eta_{t,0}, \dots, \eta_{t,s-1})$ , — равномерно распределенная последовательность точек  $s$ -мерного единичного куба  $E_s$ , и пусть по каждой серии  $(\eta_{t,0}, \dots, \eta_{t,s-1})$  построена стандартным способом траектория  $(\nu_0^{(t)}, \dots, \nu_{s-1}^{(t)})$ , где в качестве  $\xi_k$  в (8) подставлено  $\eta_{t,k}$ . Если функционал  $\Phi$  таков, что  $\bar{\Phi} = \Phi(\nu_0, \dots, \nu_{s-1})$ , то

$$\frac{1}{N} \sum_{t=1}^N \Phi(\nu_0^{(t)}, \dots, \nu_{s-1}^{(t)}) \xrightarrow{N \rightarrow \infty} \bar{\Phi}. \quad (9)$$

3. Таким образом, здесь для вычисления среднего приходится моделировать много траекторий фиксированной длины  $s$ . Задачи такого типа весьма многочисленны. В то же время, например, в статистической физике встречаются задачи совершенно иного типа. В них требуется вычислить среднее  $\sum_{i=1}^n q_i f(\omega^{(i)})$  по неизвестному стационарному распределению  $(q_i)$  цепи с данной матрицей  $(p_{ij})$  вероятностей перехода (см. [1, 2]). Для этого используется следующее предложение (ср. [11]).

ЛЕММА 2 (усиленный закон больших чисел для цепи Маркова). Если  $(q_i)_{i=1}^n$  — единственное стационарное распределение вероятностей стохастической матрицы  $(p_{ij})$ , то, каково бы ни было начальное распределение и функция  $f(\omega)$ ,

$$\mathbf{P} \left\{ \frac{1}{N} \sum_{k=1}^N f(\nu_k) \xrightarrow{N \rightarrow \infty} \sum_{i=1}^n q_i f(\omega^{(i)}) \right\} = 1. \quad (10)$$

Таким образом, в задачах этого типа для вычисления искомого среднего надо моделировать единственную траекторию, но зато очень большой длины. Возникает

Задача. Описать ограничения, которым должна удовлетворять псевдослучайная последовательность чисел  $\{\xi_k\}$ , чтобы, какова бы ни была стохастическая матрица  $(p_{ij})$  вероятностей переходов с единственным стационарным распределением  $(q_j)$  (такие матрицы мы будем называть *эргодическими*) и каково бы ни было начальное распределение, при моделировании этой цепи с помощью  $\{\xi_k\}$  по стандартной формуле (8) получалась эргодическая траектория  $\{\nu_k\}$ , т.е. для любой функции  $f(\omega)$

$$\frac{1}{N} \sum_{k=1}^N f(\nu_k) \xrightarrow{N \rightarrow \infty} \sum_{i=1}^n q_i f(\omega^{(i)}). \quad (11)$$

В настоящей работе мы решаем эту задачу и показываем, что обычные проверки на равномерную распределенность, рекомендуемые в руководствах [4, 5], не обеспечивают сходимости (11) среднего по построенной траектории к среднему по фазовому пространству.

Пусть по-прежнему  $\{\nu_k\}$  — последовательность исходов  $\omega \in \Omega = \{\omega^{(1)}, \dots, \omega^{(n)}\}$ . Обозначим через  $\delta_i(\omega)$  функцию Кронекера:  $\delta_i(\omega) = 1$  при  $\omega = \omega^{(i)}$  и  $\delta_i(\omega) = 0$  при  $\omega \neq \omega^{(i)}$ . Тогда число  $T_i(N)$  исходов  $\omega^{(i)}$  среди первых  $N$  исходов  $\nu_1, \dots, \nu_N$  есть

$$T_i(N) = \sum_{k=1}^N \delta_i(\nu_k)$$

и для любой функции  $f(\omega)$

$$\frac{1}{N} \sum_{k=1}^N f(\nu_k) - \sum_{i=1}^n \frac{T_i(N)}{N} f(\omega^{(i)}).$$

**ЛЕММА 3.** Для того чтобы у любой функции  $f(\omega)$  усреднение по времени вдоль траектории  $\nu_1, \dots, \nu_k, \dots$  асимптотически совпадало со средним по пространству  $\Omega$

$$\frac{1}{N} \sum_{k=1}^N f(\nu_k) \longrightarrow \sum_{i=1}^n q_i f(\omega^{(i)}),$$

необходимо и достаточно, чтобы для всех  $i$

$$\lim_{N \rightarrow \infty} N^{-1} T_i(N) = q_i. \quad (12)$$

**4.** В большинстве работ, как правило, ограничиваются только проверкой равномерной распределенности последовательности  $\{\xi_k\}$ . Покажем, что такая проверка совершенно недостаточна даже в простейших случаях.

**ПРИМЕР 1.** Рассмотрим цепь Маркова с двумя состояниями  $\omega^{(1)}$  и  $\omega^{(2)}$ , с матрицей перехода  $(p_{ij})$ ,  $p_{ij} = 0,1$  при  $i \neq j$  и  $p_{ii} = 0,9$ . Эта цепь симметричная, поэтому ее стационарное распределение равновероятное:  $q_1 = q_2 = 0,5$ . Пусть  $\{\xi_k^{(0)}, \dots, \xi_k^{(9)}\}$  — последовательность точек, равномерно распределенных в 10-мерном единичном кубе. Построим из нее новую последовательность  $\{\zeta_l\}$  по правилу

$$\zeta_{10k+r} = 0,1(r + \xi_k^{(r)}), \quad r = 0, 1, \dots, 9.$$

Нетрудно проверить, что последовательность  $\{\xi_l\}$  равномерно распределена на отрезке  $[0, 1]$  (так как каждая подпоследовательность  $\{\xi_{10k+r}\}_{k=1}^{\infty}$  равномерно распределена на отрезке  $[(r-1)/10, r/10]$ ). Выпишем подряд первые после запятой цифры чисел  $\zeta_l$ . Образуется последовательность цифр с периодом 0123456789. Переходы из  $\omega^{(1)}$  в  $\omega^{(2)}$  происходят лишь на шагах, номера которых оканчиваются на 9 (т.е. вида  $10k-1$ ), так как только тогда  $\zeta > 0,9$ ; но после этого сразу следует обратный переход, ибо  $\zeta_{10k} < 0,1$ , и там мы находимся еще при  $l = 10k+1, \dots, 10k+8$ .

Таким образом, при движении, построенном по последовательности  $\{\zeta_l\}$ , мы будем находиться в  $\omega^{(1)}$  в 9 раз чаще, чем в  $\omega^{(2)}$ :  $N^{-1} T_1(N) \rightarrow 0,9 \neq q_1 = 0,5$ , следовательно, стационарное распределение моделируется неправильно. Более сильным требованием является равномерная распределенность последовательных пар или троек, или четверок, где комбинации берутся гусеницей (4). Последовательности, обладающие этим свойством, мы будем называть *двукратно*, *трехкратно*

и, в общем случае,  $s$ -кратно равномерно распределенными. Чтобы разобраться и здесь, рассмотрим ту же цепь, что и в примере 1, и построим последовательность  $\{\zeta_l\}$  так, чтобы первые после запятой цифры  $\beta_l$  чисел  $\zeta_l$  повторялись с периодом  $000102 \dots 091011 \dots 9899$  длиной двести. А чтобы последовательность пар  $(\zeta_l, \zeta_{l+1})$  была равномерно распределена в  $E_2$ , возьмем числовую последовательность  $\{\xi_k\}$ , у которой пары  $\{(\xi_k, \xi_{k+1})\}$  равномерно распределены в единичном квадрате,  $\xi_k = 0, \alpha_k^{(1)} \alpha_k^{(2)} \alpha_k^{(3)} \dots$ , и допишем последующие цифры  $\zeta_l$  по правилу

$$\zeta_{200k+m} = 0, \beta_{200k+m} \alpha_k^{(m)} \alpha_k^{(m+200)} \alpha_k^{(m+400)} \dots, \quad m = 1, \dots, 200.$$

Нетрудно проверить, что на длине периода мы 164 раза побываем в  $\omega^{(1)}$  и только 36 раз в  $\omega^{(2)}$ , т.е.  $\lim N^{-1} T_2(N) = 0,18$ , так что и здесь стационарное распределение моделируется неправильно. Общий случай разбирается сходным образом.

5. Следуя Н.М. Коробову (см. [3], а также [12]), станем называть числовую последовательность  $\{\xi_k\}$  *вполне равномерно распределенной*, если она  $s$ -кратно равномерно распределена при всех натуральных  $s$ , т.е. имеет следующее

Свойство В. Для любого  $s, s = 1, 2, \dots$ , построенная по  $\{\xi_k\}$  последовательность векторов  $\eta_k = (\xi_k, \xi_{k+1}, \dots, \xi_{k+s-1})$  равномерно распределена в единичном кубе  $E_s$ . Свойство вполне равномерной распределенности не является чересчур ограничительным: с вероятностью единица последовательность независимых равномерно распределенных случайных чисел будет не только равномерно распределена по Вейлю, но и вполне равномерно распределена по Коробову<sup>2)</sup>. Конкретные примеры таких последовательностей приведены в [3, 12].

Простейшими эргодическими стохастическими матрицами являются стохастические матрицы со строго положительными элементами. Для этого класса цепей Маркова наше исследование существенно упрощается.

ТЕОРЕМА 1. Пусть у матрицы  $(p_{ij})_{i,j=1}^n$  моделируемой цепи все элементы положительны. Если последовательность  $\{\xi_k\}$  вполне равномерно распределена, то сконструированная по ней согласно (8) траектория  $\{\nu_k\}$  является эргодической, т.е.

$$\lim_{N \rightarrow \infty} N^{-1} T_i(N) = q_i, \quad i = 1, \dots, n,$$

где  $(q_1, \dots, q_n)$  — единственное стационарное распределение цепи.

Доказательство. Воспользуемся тем, что при строго положительной матрице  $(p_{ij})$  положение  $\nu_k$  на “разыгрываемой” траектории не зависит от далеких предшествующих положений. Пусть  $p = \min_i p_{in}$ . Очевидно, при  $\xi > 1 - p$  положение  $\nu_k = \omega^{(n)}$ , каковы бы ни были состояния  $\nu_{k-1}, \nu_{k-2}$  и т.д., и все последующие положения определяются только значениями  $\xi_{k+1}, \xi_{k+2}, \dots$ . Пусть  $\Lambda_k^{(m)} = (\lambda_{k,0}, \lambda_{k,1}, \dots, \lambda_{k,m})$  — последовательность серий исходов  $\omega$ , где исход  $\lambda_{k,0}$  определяется по  $\xi_k$  и стационарному распределению  $(q_i)_{i=1}^n$ , а все  $\lambda_{k,t}$  при  $t \geq 1$  “разыгрываются” по матрице перехода  $(p_{ij})$  из предыдущего положения  $\lambda_{k,t-1}$

<sup>2)</sup> Отметим еще следующее. Пусть последовательность  $\{\xi_k\}$  такова, что при любом  $s$  построенная из нее по формуле (5) последовательность векторов  $\zeta_k = (\xi_{ks+1}, \dots, \xi_{ks+s})$  равномерно распределена в  $E_s$ . Тогда последовательность  $\{\xi_k\}$  обязательно вполне равномерно распределена по Коробову (см. [13]).

по  $\xi_{k+t}$  посредством правила (8). Так как  $\{\xi_k\}$  — вполне равномерно распределенная последовательность, то по лемме 1 при любом  $t$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \delta_i(\lambda_{k,m}) = q_i. \quad (13)$$

Область  $\{y_1 < 1-p, y_2 < 1-p, \dots, y_m < 1-p\}$  есть  $m$ -мерный параллелепипед  $A_m$  объема  $(1-p)^m$ . Очевидно, если  $(\xi_{k+1}, \dots, \xi_{k+m}) \notin A_m$ , то  $\lambda_{k,m} = \nu_{k+m}$ . Поэтому

$$\frac{1}{N} \left| \sum_{k=1}^N [\delta_i(\lambda_{k,m}) - \delta_i(\nu_{k+m})] \right| \leq \mu_N(A_m), \quad (14)$$

где  $\mu_N(A_m) = F_N^*(1-p, 1-p, \dots, 1-p)$  — доля первых  $N$  векторов  $(\xi_{t+1}, \dots, \xi_{t+m})$ , попавших в  $A_m$ , или, если угодно, среднее арифметическое значений характеристической функции параллелепипеда, взятых в первых  $N$  точках. В силу  $m$ -кратной равномерной распределенности  $\{\xi_k\}$

$$\lim_{N \rightarrow \infty} \mu_N(A_m) = (1-p)^m. \quad (15)$$

Наконец, нам потребуется еще одна оценка:

$$\left| \frac{1}{N} \sum_{k=1}^N \delta_i(\nu_{k+m}) - \frac{1}{N} \sum_{k=1}^N \delta_i(\nu_k) \right| < \frac{m}{N}. \quad (16)$$

Теперь мы можем сравнить  $q_i$  с  $N^{-1}T_i(N)$ :

$$\begin{aligned} |N^{-1}T_i(N) - q_i| &\leq \left| N^{-1} \sum_{k=1}^N \delta_i(\nu_k) - N^{-1} \sum_{k=1}^N \delta_i(\nu_{k+m}) \right| + \\ &+ \left| N^{-1} \sum_{k=1}^N [\delta_i(\nu_{k+m}) - \delta_i(\lambda_{k,m})] \right| + \left| N^{-1} \sum_{k=1}^N \delta_i(\lambda_{k,m}) - q_i \right|. \end{aligned}$$

Верхний предел при  $N \rightarrow \infty$  второго слагаемого в силу (14) и (15) может быть сделан сколь угодно малым при достаточно большом  $t$ . При фиксированном  $t$  пределы первого и третьего слагаемых существуют и равны нулю; см. (16) и (13). Следовательно,  $N^{-1}T_i(N) \rightarrow q_i$ .

Теорема доказана.

Теорема 1 является частным случаем следующего, более общего предложения.

**ТЕОРЕМА 2.** Если последовательность  $\{\xi_k\}$  вполне равномерно распределена, то для любой стохастической матрицы  $(p_{ij})$  с единственным стационарным распределением  $(q_i)$  стандартная конструкция (8) отображает  $\{\xi_k\}$  в эргодичную в смысле (11) траекторию  $\{\nu_k\}$ . Обратно, если для любой стохастической матрицы с единственным стационарным распределением стандартная конструкция определяет по  $\{\xi_k\}$  эргодичную траекторию, то  $\{\xi_k\}$  вполне равномерно распределено.

Доказательство обратной теоремы. Для проверки вполне равномерной распределенности  $\{\xi_k\}$  достаточно проверить при всех  $s$  сходимость  $F_N^*(z_1, \dots, z_s)$  к  $z_1 \cdot \dots \cdot z_s$  в любой точке  $z$  куба  $E_s$ . Построим следующее семейство цепей Маркова.

ПРИМЕР 2. Пусть имеются  $2s$  состояний  $\omega^{(1)}, \dots, \omega^{(s)}, \omega^{(s+1)}, \dots, \omega^{(2s)}$  со следующими вероятностями переходов: если  $1 < i \leq s$ , то  $p_{i-1,i} = z_i$  и  $p_{i-1,i+s} = 1 - z_i$ ; если  $s + 1 < i \leq 2s$ , то  $p_{i-1,i} = 1$ ; далее,  $p_{s,1} = p_{2s,1} = z_1$ ,  $p_{s,s+1} = p_{2s,s+1} = 1 - z_1$ , а все остальные переходные вероятности равны нулю.

Все состояния цепи образуют один класс сообщающихся состояний, разбивающийся на  $s$  подклассов  $\Omega_i = (\omega^{(i)}, \omega^{(s+i)})$ ,  $i = 1, \dots, s$ . Поэтому (ср. [14]) стационарное распределение единственно и, как нетрудно подсчитать,  $q_s = s^{-1} z_1 \cdot \dots \cdot z_s$ . Какова бы ни была последовательность  $\{\xi_k\}$ , траектория все время переходит из одного подкласса в следующий в циклическом порядке. Если  $\nu_0 = \omega^{(2s-l)}$ ,  $l = 0, \dots, s-1$ , то в  $\omega^{(s)}$  мы можем попадать только на шагах вида  $k = ms + l$ ,  $m = 1, 2, \dots$ , причем будем там в том и только том случае, если  $(\xi_{k-s+1}, \dots, \xi_k) \in B$ , где  $B$  — параллелепипед  $\{0 \leq x_1 < z_1, \dots, 0 \leq x_s < z_s\}$ . Обозначим  $\chi_B(z_1, \dots, z_s)$  характеристическую функцию параллелепипеда  $B$ . Так как сконструированная по  $\{\xi_k\}$  траектория предполагается эргодичной, то при всех  $l$ ,  $1 \leq l \leq s$ ,

$$N^{-1}T_s(N) = \frac{1}{N} \sum_{m=1}^{[N_s-1]} \chi_B(\xi_{ms-l+1}, \dots, \xi_{ms-l+s}) \xrightarrow{N \rightarrow \infty} \frac{1}{s} z_1 \cdot \dots \cdot z_s. \quad (17)$$

Суммируя (17) по всем начальным положениям  $\omega^{(1)}, \dots, \omega^{(2s)}$  и переставляя порядок суммирования по  $m$  и  $l$ , получаем

$$F_N^*(z_1, \dots, z_s) = \frac{1}{N} \sum_{k=s}^N \chi_B(\xi_{k-s+1}, \dots, \xi_k) \xrightarrow{N \rightarrow \infty} z_1 \cdot \dots \cdot z_s,$$

что и требовалось.

Доказательство прямой теоремы. Из усиленного закона больших чисел (10) вытекает, что для эргодической цепи Маркова

$$\mathbf{P}\{|N^{-1}T_i(N) - q_i| > \varepsilon\} \xrightarrow{N \rightarrow \infty} 0 \quad (18)$$

при  $N \rightarrow \infty$  для любого  $i$ , любого начального  $\omega^{(j)}$  и любого  $\varepsilon > 0$ . Независимое доказательство этого последнего утверждения (называемого просто законом больших чисел) см. в [15]. Обозначим через  $C_{ji}^{(N)}(\varepsilon)$  множество траекторий  $\lambda_1, \lambda_2, \dots, \lambda_N$ , начинающихся в  $\lambda_0 = \omega^{(j)}$ , у которых  $|N^{-1}T_i(N) - q_i| > \varepsilon$ , и, согласно (18), выберем такое  $N = s$ , чтобы

$$\max_j \mathbf{P}\{C_{ji}^{(s)}(\varepsilon) | \lambda_0 = \omega^{(j)}\} < \varepsilon n^{-1}.$$

Пусть  $D_{ji}^{(s)}(\varepsilon)$  — множество всех векторов  $(x_1, \dots, x_s) \in E_s$  таких, что построенная по ним стандартным способом (8), начиная с  $\lambda_0 = \omega^{(j)}$ , траектория  $\lambda_1, \dots, \lambda_s$

принадлежит  $C_{ji}^{(s)}(\varepsilon)$ , т.е.  $D_{ji}^{(s)}(\varepsilon)$  — полный прообраз множества  $C_{ji}^{(s)}(\varepsilon)$ . Отображающая вектор-функция  $\varphi(x_1, \dots, x_s) = (\lambda_1, \dots, \lambda_s)$  принимает конечное число значений, постоянных на каждом из параллелепипедов, определяемых последовательностью условий (8). Поэтому множество  $D_{ji}^{(s)}(\varepsilon)$  измеримо по Жордану и его объем, равный  $\mathbf{P}\{C_{ji}^{(s)}(\varepsilon) | \lambda_0 = \omega^{(j)}\}$ , не превосходит  $\varepsilon n^{-1}$ . Объединение  $D_i^{(s)}(\varepsilon) = \bigcup^n D_{ji}^{(s)}(\varepsilon)$  также измеримо по Жордану и может быть разбито на параллелепипеды; при этом объем  $V_i^{(s)}(\varepsilon)$  множества  $D_i^{(s)}(\varepsilon)$  не превосходит  $\varepsilon$ .

Пусть теперь  $\{\nu_k\}$  — траектория, сконструированная по  $\{\xi_k\}$ . Обозначим через  $Q_i^{(s)}(N, \varepsilon)$  множество тех точек  $\{\xi_k, \dots, \xi_{k+s-1}\}$ ,  $k = 1, \dots, N$ , которые принадлежат  $D_i^{(s)}(\varepsilon)$ , и через  $d_i^{(s)}(N, \varepsilon)$  — их число. Пусть

$$m_i^{(s)}(k) = \frac{1}{s} \sum_{t=0}^{s-1} \delta_i(\nu_{k+t}).$$

Для любого  $k \notin Q_i^{(s)}(N, \varepsilon)$  обязательно  $|m_i^{(s)}(k) - q_i| \leq \varepsilon$ . Заметив это, разобьем  $T_i(N)$  на сумму трех слагаемых:

$$T_i(N) = \sum_{k \notin Q(N)} m_i^{(s)}(k) + \sum_{k \in Q(N)} m_i^{(s)}(k) + \sum_{t=1}^s \frac{t}{s} [\delta_i(\nu_t) - \delta_i(\nu_{N+s+1-t})]. \quad (19)$$

Тогда для разности  $|N^{-1}T_i(N) - q_i|$  можно написать следующую оценку:

$$|N^{-1}T_i(N) - q_i| \leq \varepsilon [1 - N^{-1}d_i^{(s)}(N, \varepsilon)] + N^{-1}d_i^{(s)}(N, \varepsilon) + N^{-1}s.$$

В силу  $s$ -кратной равномерной распределенности  $\{\xi_k\}$

$$N^{-1}d_i^{(s)}(N, \varepsilon) \xrightarrow{N \rightarrow \infty} V_i^{(s)}(\varepsilon) \leq \varepsilon$$

и  $N^{-1}s \rightarrow 0$  при фиксированном  $s$ . Поэтому, начиная с некоторого  $N_0 = N_0(s, \varepsilon)$ ,

$$|N^{-1}T_i(N) - q_i| \leq 4\varepsilon.$$

Таким образом,  $N^{-1}T_i(N) \rightarrow q_i$  при  $N \rightarrow \infty$  для каждого  $i$ , что и требовалось доказать.

ПРИМЕЧАНИЕ. Кроме средних по стационарному распределению  $(q_i)$ , могут быть интересны средние

$$\sum_{i,j} g(\omega^{(i)}, \omega^{(j)}) q_i p_{ij}$$

по стационарному распределению переходов, а также средние величины, определяемые несколькими последовательными состояниями. Так как методы работы [15] позволяют доказать закон больших чисел и для переходов  $(\omega^{(i)}, \omega^{(j)})$ , и для появления более длинных комбинаций, то доказательство теоремы 2 переносится на этот случай без особых изменений. Поэтому вполне равномерная распределенность псевдослучайной последовательности обеспечивает асимптотически правильное моделирование и для вычисления таких более сложных характеристик, так сказать, вполне эргодичность построенной траектории.

**6.** Получим теперь из теоремы 2 ряд следствий, представляющих самостоятельный интерес.

Доказательство теоремы 2 использует лишь простой закон больших чисел для эргодических цепей Маркова (18), а не усиленный закон (10). Более того, этот закон (10) следует из прямого утверждения теоремы 2 и из сделанного в начале п. 5 замечания о том, что последовательность независимых равномерно распределенных на  $[0, 1]$  случайных чисел с вероятностью единица вполне равномерно распределена (тривиально вытекающего в свою очередь из классического усиленного закона больших чисел Бореля для независимых событий; ср., впрочем, [16]). Таким образом, мы получили формально новое доказательство усиленного закона больших чисел для эргодических цепей Маркова.

При доказательстве обратного утверждения теоремы 2 мы доказали сначала для любого натурального  $s$ , любого  $l \leq s$  и любого  $z \in E_s$  необходимость выполнения более ограничительного соотношения (17) и уже из него вывели утверждение теоремы. Поэтому для того чтобы последовательность  $\{\xi_k\}$  была “хорошей”, т.е. пригодной для моделирования стационарных распределений эргодических цепей Маркова, достаточно, чтобы  $\{\xi_k\}$  была вполне равномерно распределенной, и необходимо, чтобы она обладала следующим формально более сильным свойством А.

**СВОЙСТВО А.** Для любого  $s = 1, 2, \dots$  и любого  $l, 1 \leq l \leq s$ , последовательность векторов  $\{(\xi_{rs-l+1}, \dots, \xi_{rs-l+s})\}_{r=1}^{\infty}$  равномерно распределена в  $E_s$ .

Но раз так, то свойство А и свойство В вполне равномерной распределенности последовательности эквивалентны друг другу. Рассмотрим еще последовательности  $\{\xi_k\}$  с формально более слабым, чем А, свойством Б.

**СВОЙСТВО Б.** Для любого  $s = 1, 2, \dots$  и некоторого  $l = l(s), 1 \leq l \leq s$ , последовательность векторов  $\{(\xi_{rs-l+1}, \dots, \xi_{rs-l+s})\}_{r=1}^{\infty}$  равномерно распределена в  $E_s$ .

Нетрудно показать, что прямое утверждение теоремы 2 сохранится, если в его формулировке заменить свойство В на свойство Б. Для этого надо лишь слегка изменить его доказательство, выражая величину  $T_i(N)$  не через все средние  $m_i^{(s)}, k = 1, \dots, N$ , как в (19), а лишь через средние  $m_i^{(s)}(k_r)$  по непересекающимся отрезкам  $k_r = rs - l + 1, k_r + 1, \dots, k_r + s = r(s + 1) - l$  длины  $s$  ряда  $1, \dots, N$ . Поэтому из Б также следует А. Тем самым доказана

**ЛЕММА 4.** Свойства А, Б и В числовых последовательностей  $\{\xi_k\}_{k=1}^{\infty}, 0 \leq \xi_k \leq 1$ , эквивалентны друг другу, и любое из них может быть принято за определение вполне равномерной распределенности по Коробову.

Разумеется, можно доказать (см. [13]) эквивалентность свойств А и Б непосредственно из определения. Прямой же вывод свойств А из свойства В использует закон больших чисел (правда, только в форме Бернулли; см. [14]) для независимых событий.

В работе [17] был впервые указан способ построения вполне эргодических траекторий марковской цепи с двумя состояниями, которые были названы нормальными по Маркову последовательностями двоичных знаков. Как следует из примечания к теореме 2, стандартная конструкция (8) задает конструкцию нормальных по Маркову последовательностей из  $n$  знаков с переходной матрицей  $(p_{ij})_{i,j=1}^n$  (ср. также [18]).

7. Было бы очень важно описать условия, позволяющие моделировать стационарное распределение любой эргодической марковской цепи уже не обязательно с конечным числом состояний. Пока не ясно, достаточно ли здесь вполне равномерной распределенности применяемых для розыгрыша псевдослучайных чисел. Во всяком случае, доказательство теоремы 2 существенно использует конечность числа состояний и непосредственно обобщено быть не может. Однако можно модифицировать доказательство более простой теоремы 1 и указать класс марковских переходных операторов, где для моделирования хватает вполне равномерной распределенности. Пусть множество состояний — числовая прямая  $-\infty < \omega < \infty$ , и пусть переходное распределение вероятностей задается условной функцией распределения  $F_y(x) = P\{\nu_k < x | \nu_{k-1} = y\}$ . Тогда неравенство (8) переходит в уравнение  $F_y(x) = \xi_k$ , где  $\nu_{k-1} = y$ , и корень  $x = F_y^{-1}(\xi_k)$  уравнения принимается за  $\nu_k$  (ср. [4]). Назовем условную переходную функцию  $F_y(x)$  *сжимающей*, если при любом  $\xi$  справедливо неравенство  $|F_y^{-1}(\xi) - F_z^{-1}(\xi)| < c|y - z|$ , где  $c < 1$ . Тогда разыгрываемая траектория также экспоненциально “забывает” свое прошлое, и идея доказательства теоремы 1 может быть проведена и здесь.

ПРИМЕР 3. Пусть множество состояний  $\Omega = [0 \leq \omega \leq 1]$  — единичный отрезок, плотность вероятности перехода  $p(x|\omega) = (1-x)(1-\omega) + x\omega$ . Соответствующая условная функция распределения перехода есть  $F_\omega(x) = (2\omega-1)x^2 + 2(1-\omega)x$ . Дифференцируя соотношение  $F_\omega(x) = \xi$  при постоянном  $\xi$ , получаем

$$\frac{\partial x}{\partial \omega} = \frac{x(1-x)}{(1-x)(1-\omega) + x\omega} \leq 1,$$

причем равенство возможно лишь в точках  $\omega = 0$ ,  $x = 1$  и  $\omega = 1$ ,  $x = 0$ . Поэтому при вполне равномерно распределенной последовательности  $\{\xi_k\}$  далекое прошлое по большей части (так сказать, с большой вероятностью) почти не влияет на моделируемое настоящее, и траектория будет эргодичной.

ПРИМЕЧАНИЕ. Настоящая работа, за исключением теоремы 2 и п. 6, была доложена на семинаре по теоретико-числовым методам в приближенном анализе в Математическом институте АН СССР.

Автор считает приятным долгом выразить Н.М. Коробову искреннюю признательность за ценную дискуссию.

## СПИСОК ЛИТЕРАТУРЫ

1. Фишер И.З. Статистическая теория жидкостей. — М.: Физматгиз, 1961.
2. Metropolis N., Rosenbluth M.N., Rosenbluth A.W., Teller A., Teller E. Equation of state calculation by fast computing machines // J. Chem. Phys. 1953. V. 21, № 6. P. 1087–1092.
3. Коробов Н. М. О некоторых вопросах равномерного распределения // Изв. АН СССР. Сер. матем. 1950. Т. 14, № 3. С. 215–238.
4. Бусленко Н.П. и др. Метод статистических испытаний (метод Монте-Карло). — М.: Физматгиз, 1962.
5. Голенко Д.И. Моделирование и статистический анализ псевдослучайных чисел на электронных вычислительных машинах. — М.: Наука, 1965.
6. Колмогоров А.Н. Основные понятия теории вероятностей. — М.-Л.: ОНТИ, 1936.
7. Полюа Г., Сегё Г. Задачи и теоремы из анализа. Ч. I. — М.: Гостехиздат, 1956.
8. Weyl H. Über die Gleichverteilung von Zahlen mod Eins // Math. Ann. 1916. V. 77. P. 313–352.

9. *Соболев И.М.* Точная оценка погрешности многомерных квадратурных формул для функций классов  $\widetilde{W}_1$  и  $\widetilde{H}_1$  // ЖВМиМФ. 1961. Т. 1, № 2, С. 208–215.
10. *Гантмахер Ф.Р.* Теория матриц. — М.: Наука, 1966.
11. *Сарымсаков Т.А.* Основы теории процессов Маркова. — М.: Гостехиздат, 1954.
12. *Коробов Н.М.* О вполне равномерном распределении и совместно нормальных числах // Изв. АН СССР. Сер. матем. 1956 Т. 20, № 6. С. 649–660.
13. *Коробов Н.М.* Об арифметических свойствах показательных функций: Дис. ... д-ра физ.-мат. наук. — М.: МИ АН СССР, 1953.
14. *Лозв М.* Теория вероятностей. — М.: ИЛ, 1962.
15. *Колмогоров А.Н.* Локальная предельная теорема для классических цепей Маркова // Изв. АН СССР. Сер. матем. 1949. Т. 13, № 4. С. 281–300.
16. *Постников А.Г.* Усиленный закон больших чисел для выборки из равномерно распределенной величины // Изв. АН СССР. Сер. матем. 1958. Т. 22, № 3. С. 433–438.
17. *Постников А.Г., Пятецкий И.И.* Нормальная по Маркову последовательность знаков и нормальная цепная дробь // Изв. АН СССР. Сер. матем. 1957. Т. 21, № 5. С. 729–746.
18. *Шахов Ю.Н.* Имитация простейших марковских процессов // Изв. АН СССР. Сер. матем. 1959. Т. 23, № 5. С. 815–822.

## РАСПРОСТРАНЕНИЕ В ВОЗДУХЕ $\gamma$ -ИЗЛУЧЕНИЯ ИЗ МГНОВЕННОГО ТОЧЕЧНОГО ИСТОЧНИКА <sup>\*)</sup>

Для расчета методом Монте-Карло задачи о распространении в бесконечной воздушной среде импульса  $\gamma$ -излучения бесконечно малой длительности ( $\delta$ -импульс) из точечного источника мы брали начальную энергию  $\gamma$ -излучения, равную 1 МэВ, и плотность воздуха  $1,29 \cdot 10^{23}$  г/см<sup>3</sup>. Необходимые при расчетах значения полного коэффициента взаимодействия  $\gamma$ -излучения с воздухом и величины сечений образования пар комптоновского и фотоэлектрического эффектов были взяты из работ [1, 2].

Для решения задачи точно моделировалось движение  $\gamma$ -кванта, испущенного источником. Направление движения кванта после каждого столкновения разыгрывалось по индикатрисе Клейна–Нишины. Для быстрого решения получающегося трансцендентного уравнения была разработана специальная методика. Поглощение квантов средой не разыгрывалось, а учитывалось весовым множителем [3], поэтому по каждой траектории двигался пакет квантов. Траектория пакета прослеживалась вплоть до веса, равного  $10^{-4}$  начального веса. Согласно работе [4] после каждого столкновения подсчитывалась доля пакета квантов, которые, двигаясь без столкновений по разыгранному направлению, достигли бы сферы радиусом  $R$ . Такая регистрация  $\gamma$ -излучения проводилась сразу для

<sup>\*)</sup> Атомная энергия. 1961. Т. 10, вып. 5. С. 493–500 (совм. с О.И. Лейпунским, А.С. Стрелковым, А.С. Фроловым).

нескольких радиусов  $R_k$ . Предлагаемая методика позволила по сравнительно небольшому количеству траекторий получить приемлемое описание процесса распространения  $\gamma$ -импульса вплоть до расстояний в восемь-десять длин свободного пробега начального излучения.

Розыгрыш случайных параметров производили по последовательности псевдослучайных чисел типа чисел, предложенных в работе [5], генерировавшихся программой. В излагаемом варианте задачи число испытаний составляло 1500. Для оценки точности дополнительно определяли количество  $\gamma$ -квантов, проходящих сквозь каждую сферу радиусом  $R_k$ , а также среднюю ошибку этой величины.

В результате расчета находили величины  $\Phi_{kjim}$ , представляющие для данного расстояния  $R_k$  значения энергии, которая переносится за время  $t_j - t_{j+1}$  через единичную площадку, перпендикулярную к потоку,  $\gamma$ -квантами с энергией  $E_i - E_{i+1}$ , имеющими угол  $\theta_m - \theta_{m+1}$  между направлением движения и радиусом-вектором точки наблюдения. За нуль времени для каждой точки наблюдения принималось время прихода в нее прямого излучения. На основе величин  $\Phi_{kjim}$  были найдены соответствующие величины временных энергетических угловых интенсивностей излучения  $I_{kjim}^0 = \Phi_{kjim} / (\Delta t_j \Delta E_i \Delta \Omega_m)$ . При расчете значения  $R_k$ ,  $t_j$ ,  $E_i$  и  $\theta_m$  были приняты: 1)  $R_k = 250, 500, 1000$  м, что соответствует значению  $\mu_0 R_k$ , равному 2,03; 4,06; 8,12 длины свободного пробега; 2)  $t_j = 0; 0,125; 0,250; 0,500; 1,00; 1,50; 2,00; 3,00; 4,00; \infty$  мкс; 3)  $E_i = 0; 0,0625; 0,125; 0,250; 0,500; 1,00; 2,00$  МэВ; 4)  $\theta_m = 0; 10; 40; 90; 180^\circ$ . Остальные величины, относящиеся к распространению  $\gamma$ -импульса, находили суммированием величины  $\Phi_{kjim}$  или  $I_{kjim}^0$  по интересующим нас переменным. Приводимые далее результаты расчетов относятся либо к одному испущенному в импульсе  $\gamma$ -кванту (для временных зависимостей), либо к источнику мощностью 1 квант/с (зависимости, полученные для постоянного источника).

Таблица 1

Сравнение факторов накопления

$R_k$ , м ( $\mu_0 R_k$ )	250 (2,03)	500 (4,06)	1000 (8,12)
Расчет методом Монте-Карло	3,69	7,57	21,8
Расчет по методу работы [6]	3,6	7,5	18,6

С целью проверки правильности расчета путем интегрирования  $I_{kjim}^0$  находили значения факторов накопления (табл. 1) и дифференциальные энергетические спектры интенсивности рассеянного  $\gamma$ -излучения (рис. 1), которые сравнивались с результатами работы [6]. Сравнение показало довольно хорошее совпадение результатов, что в свою очередь подтверждает правильность произведенного расчета.

Рассмотрим теперь данные, полученные для  $\delta$ -источника. На рис. 2 а, б, в представлены зависимости интенсивности рассеянного  $\gamma$ -излучения от времени. За единицу интенсивности принималась интенсивность за время 0–0,125 мкс; абсолютное значение интенсивности за этот интервал времени в расчете на всю поверхность сферы ( $4\pi R^2 I_{k1}^0$ ) для  $R = 250$  м равно 1,43 МэВ/мкс, для  $R = 500$  м — 0,41 МэВ/мкс и для  $R = 1000$  м — 0,0088 МэВ/мкс. Интенсивность первичного

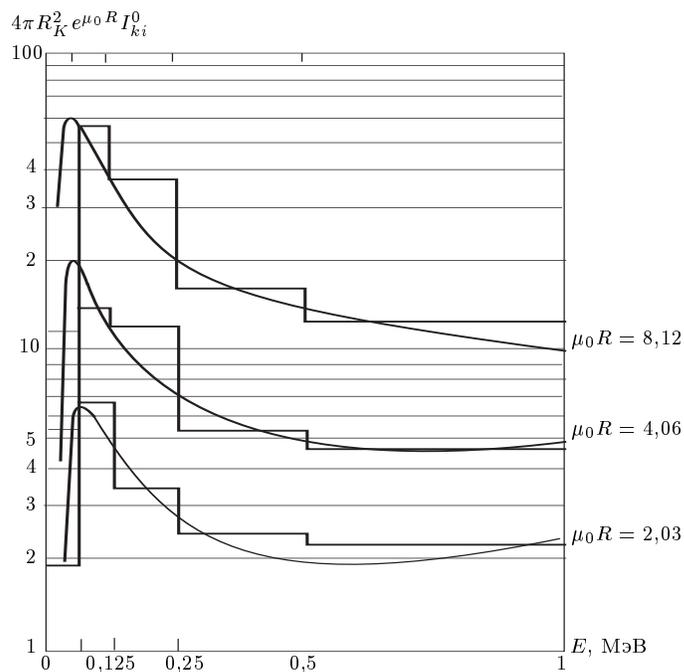


Рис. 1. Дифференциальные энергетические спектры интенсивности рассеянного  $\gamma$ -излучения с начальной энергией 1 МэВ постоянно действующего источника  $\gamma$ -квантов. Плавные кривые — данные работы [6]

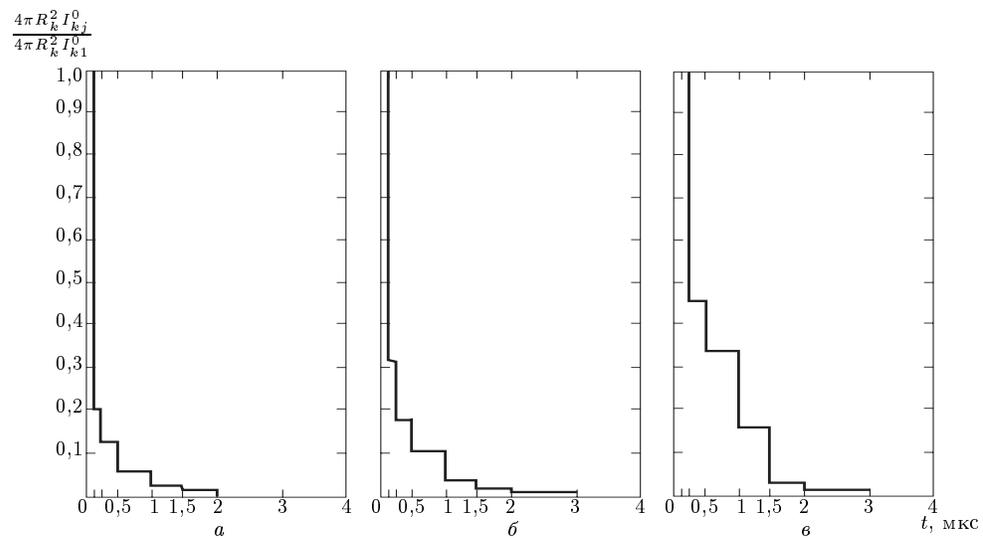


Рис. 2. Зависимости интенсивности рассеянного  $\gamma$ -излучения от времени для трех расстояний точки наблюдения от источника (в метрах): а — 250; б — 500; в — 1000

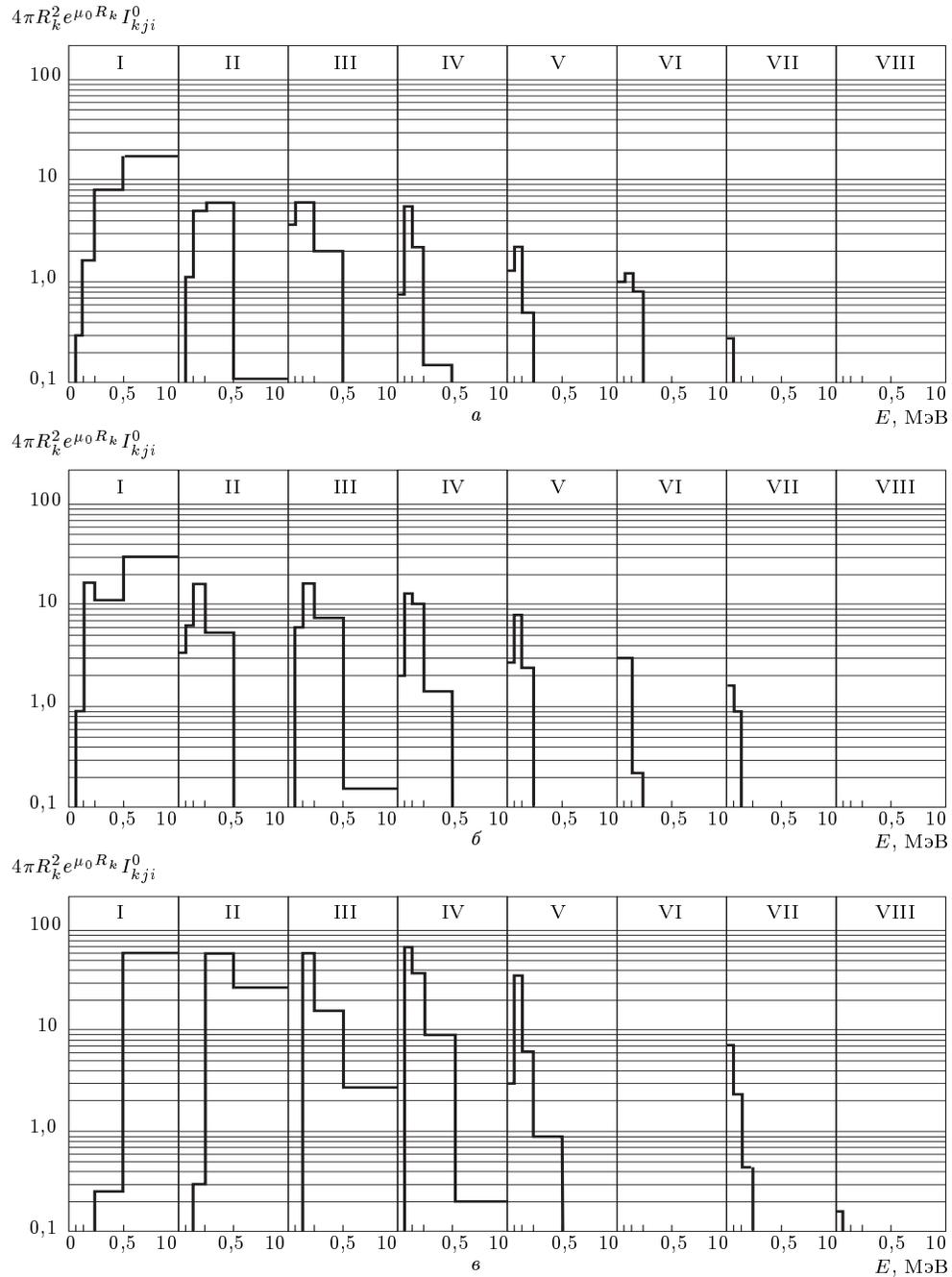


Рис. 3. Временные энергетические спектры интенсивности рассеянного  $\gamma$ -излучения для трех расстояний точки наблюдения от источника (в метрах): а — 250; б — 500; в — 1000 (I —  $0 < t < 0,125$ , II —  $0,125 < t < 0,25$ , III —  $0,25 < t < 0,5$ , IV —  $0,5 < t < 1$ , V —  $1 < t < 1,5$ , VI —  $1,5 < t < 2$ , VII —  $2 < t < 3$ , VIII —  $3 < t < 4$ )

излучения на рис. 2 не указана, так как она представлялась бы  $\delta$ -функцией с бесконечной амплитудой. Сопоставление рис. 2, *a*, *б*, *в* показывает, что по мере удаления точки наблюдения импульс растягивается (по крайней мере до расстояний в  $\mu_0 R = 8,12$ ). Время спада интенсивности до 0,1 начальной величины для расстояний 250, 500 и 1000 м равно соответственно 0,5, 1,0 и 1,5 мкс. Такое уширение импульса можно объяснить, если учесть, что на больших расстояниях в спектре излучения начинают преобладать мягкие многократно рассеянные  $\gamma$ -кванты, которые проделали большой путь и приходят в точку наблюдения с большим запозданием (относительно времени прихода прямого излучения). Это же подтверждается полученными в расчете временными энергетическими спектрами интенсивности рассеянного  $\gamma$ -излучения (рис. 3, *a*, *б*, *в*). Величина  $4\pi R_k^2 e^{\mu_0 R_k} I_{kji}^0$  пропорциональна интенсивности излучения с энергией  $E_i - E_{i+1}$  (в расчете на единичный интервал энергии) в момент времени  $t_j - t_{j+1}$ . Значения интервалов времени приведены в верхней части рисунка. Из рис. 3 видно, что для данного интервала времени форма спектров на расстояниях, больших 250 м, приблизительно сохраняется. Как и следовало ожидать, с увеличением времени в спектрах начинает преобладать мягкое многократно рассеянное излучение. Начиная с  $t = 1$  мкс, в спектрах практически не остается излучения с энергией более 0,25 МэВ. Изменение средней по спектру энергии рассеянного  $\gamma$ -излучения со временем приведено на рис. 4, *a*, *б*, *в*. На рис. 4, *a* плавной кривой показано изме-

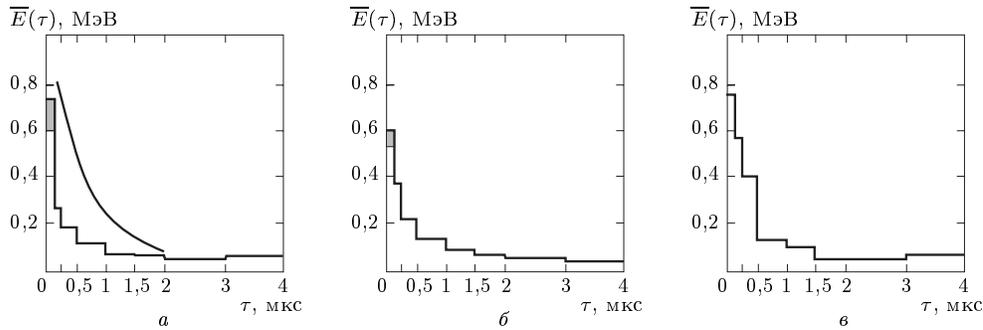


Рис. 4. Зависимость средней энергии рассеянного  $\gamma$ -излучения от времени при трех расстояниях точки наблюдения от источника (в метрах): *a* — 250; *б* — 500; *в* — 1000. Заштрихованные части гистограммы показывают увеличение жесткости при учете прямого излучения

нение энергии  $\gamma$ -излучения с начальной энергией 1 МэВ, вычисленное по средним углам комптоновского рассеяния и средним временам жизни между последовательными рассеяниями:

$$E_{n+1} = E_n \cos \theta(E_n), \quad \tau_{n+1} = \frac{\lambda_n}{c} = \frac{1}{\sigma_k(E_n) \rho_{\text{возд}} c}.$$

Сравнение показывает, что при всех временах средняя по спектру жесткость излучения меньше, чем средняя жесткость излучения, вычисленная по значениям  $E_{n+1}$  и  $\tau_{n+1}$ . Кроме того, из рис. 4 видно, что начиная с  $t = 1-1,5$  мкс средняя жесткость излучения практически не изменяется, оставаясь равной 50–60 кэВ.

Зависимость интенсивности рассеянного  $\gamma$ -излучения от времени для различных телесных углов приведена на рис. 5, *a*, *б*, *в*. На этом рисунке по оси ординат отложено отношение интенсивности в данном телесном угле (определяе-

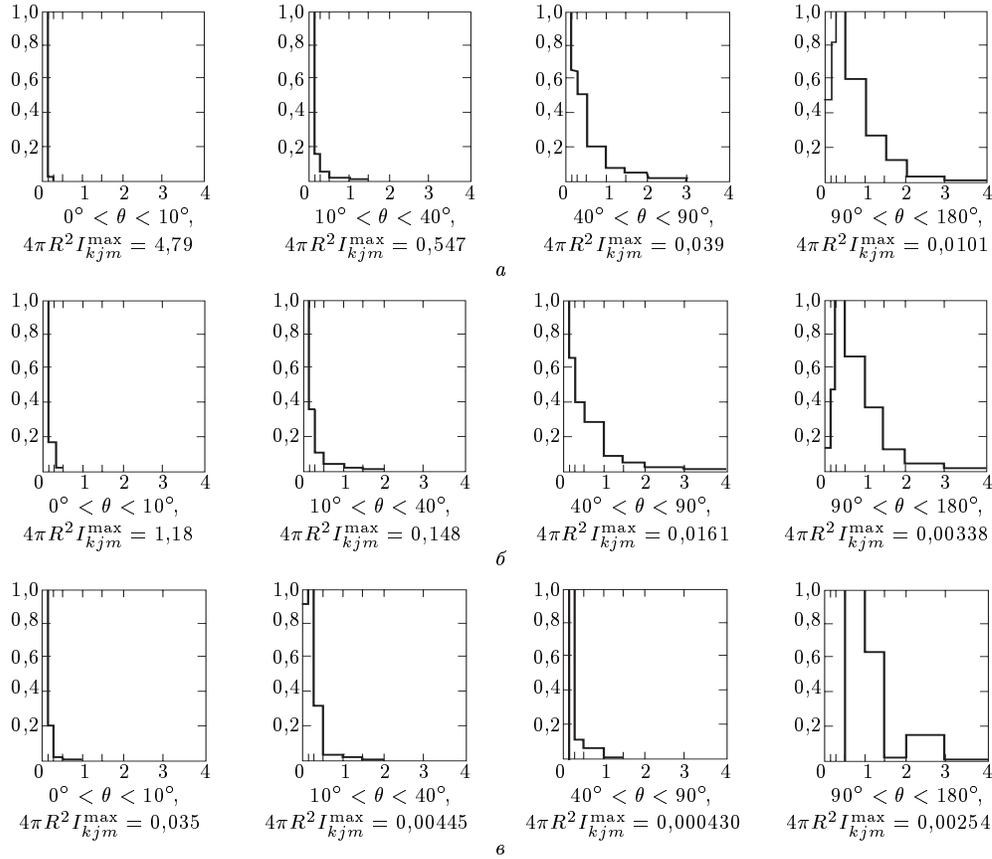


Рис. 5. Зависимость интенсивности рассеянного  $\gamma$ -излучения от времени для различных телесных углов и трех расстояний точки наблюдения от источника (в метрах, по оси абсцисс —  $t$  мкс, по оси ординат — значение величины  $I_{kjm}/I_{kjm}^{\max}$ ):  $a$  — 250;  $b$  — 500;  $в$  — 1000

мом значениями  $\theta_m < \theta < \theta_{m+1}$ ) для данного интервала времени, рассчитанной на единицу телесного угла, к максимальной интенсивности, наблюдаемой в данном телесном угле. На рисунке указаны значения максимальной интенсивности каждого из интервалов углов. Сравнивая временной ход интенсивности в определенном интервале углов  $\theta$  (или, что то же самое, в определенном телесном угле) на различных расстояниях от источника, приходим к выводу, что спад интенсивности в телесных углах, определяемых  $\theta < 90^\circ$ , затягивается с увеличением расстояния. Временной ход интенсивности излучения из заднего полупространства обнаруживает максимум, который сдвигается в сторону больших времен с увеличением расстояния от точки наблюдения.

Относительно временных энергетических угловых спектров интенсивности рассеянного излучения можно заметить следующее. С увеличением расстояния соответствующие интенсивности уменьшаются по абсолютному значению. Для данных интервалов времени и телесного угла спектры приблизительно подобны при различных  $R_k$ . С увеличением времени и  $\theta$  для данного  $R_k$  спектры смягчаются.

Остановимся теперь на данных о поле постоянно действующего источника, которые могут быть получены из решения задачи о  $\delta$ -источнике. Две характеристики поля постоянного источника — факторы накопления и дифференциальные энергетические спектры интенсивности рассеянного  $\gamma$ -излучения — уже были рассмотрены выше. Другими двумя важными характеристиками, о которых в настоящее время имеется сравнительно мало сведений, являются угловое распределение интенсивности и угловые дифференциальные спектры интенсивности рассеянного излучения.

Данные об угловом распределении интенсивности  $\gamma$ -излучения на различных расстояниях от источника приведены в табл. 2 и в табл. 3.

Таблица 2  
Угловое распределение интенсивности рассеянного  $\gamma$ -излучения

Угол $\theta$ , градусы	Расстояние от источника, м		
	250 ( $\mu_0 R = 2,03$ )	500 ( $\mu_0 R = 4,06$ )	1000 ( $\mu_0 R = 8,12$ )
0 – 10	16,9	13,5	10,4
10 – 40	37,1	37,7	32,9
40 – 90	27,8	32,3	28,9
90 – 180	17,6	16,5	27,7

Таблица 3  
Угловое распределение полной интенсивности  $\gamma$ -излучения

Угол $\theta$ , градусы	Расстояние от источника, м		
	250 ( $\mu_0 R = 2,03$ )	500 ( $\mu_0 R = 4,06$ )	1000 ( $\mu_0 R = 8,12$ )
0 – 10	39,3	24,9	14,5
10 – 40	27,1	32,7	31,4
40 – 90	20,7	28,0	27,6
90 – 180	12,8	14,3	26,5

В табл. 2 приводится доля (в процентах) интенсивности рассеянного  $\gamma$ -излучения, проходящего из данного телесного угла, по отношению ко всей интенсивности рассеянного излучения в точке наблюдения  $\left( \frac{I_{\text{рас}}^0(\Omega)}{I_{\text{рас}}^0} 100\% \right)$ ; в табл. 3 приводятся те же данные, относящиеся к полной интенсивности излучения (включая прямое). Рассматриваемые телесные углы  $\Omega = 2\pi(\cos \theta_m - \cos \theta_{m+1})$  определяются значениями  $\theta$ , приводимыми в первых колонках таблиц. Полученное в расчете угловое распределение интенсивности рассеянного излучения можно сравнить с данными, имеющимися в литературе. Исходя из измеренного в работе [7] углового распределения дозы рассеянного излучения источника  $\text{Co}^{60}$  в воде, можно вычислить процент дозы рассеянного излучения, идущего из телесных углов, рассматриваемых в настоящей работе, по отношению к полной дозе рассеянного излучения. Для  $\mu_0 R = 2$  такой подсчет дает следующие значения: при  $\theta$ , изменяющемся от 0 до 10°, доля составляет 29%; при 10–40° — 38%; при 40–90° — 25% и при 90–180° — 8,7%.

Если принять во внимание некоторые различия в начальной энергии  $\gamma$ -излучения, возмущающие действие защитных конусов, а также ход коэффициента

электронного преобразования в зависимости от энергии излучения, то можно считать, что совпадение наших результатов с результатами работы [7] является вполне удовлетворительным<sup>1)</sup>.

Угловые дифференциальные спектры интенсивности рассеянного  $\gamma$ -излучения постоянно действующего источника приведены на рис. 6. Сопоставляя

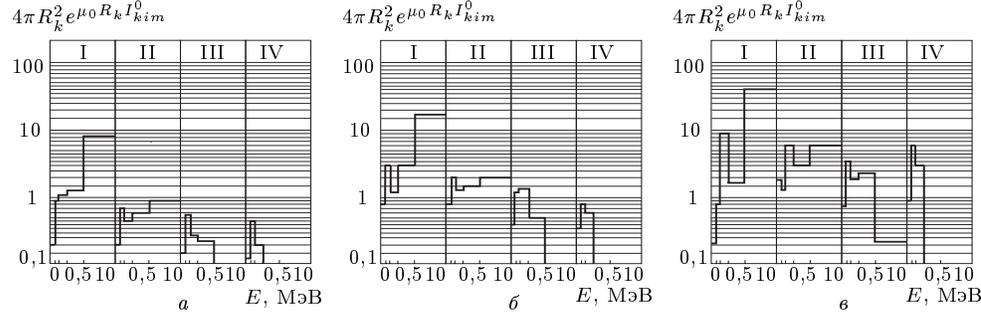


Рис. 6. Угловые дифференциальные спектры интенсивности рассеянного  $\gamma$ -излучения с энергией 1 МэВ постоянно действующего источника  $\gamma$ -квантов для трех расстояний точки наблюдения (в метрах): а — 250 ( $\mu_0 R = 2,03$ ); б — 500 ( $\mu_0 R = 4,06$ ); в — 100 ( $\mu_0 R = 8,12$ ) (I —  $0^\circ < \theta < 10^\circ$ , II —  $10^\circ < \theta < 40^\circ$ , III —  $40^\circ < \theta < 90^\circ$ , IV —  $90^\circ < \theta < 180^\circ$ )

спектры на различных расстояниях от источника, можно заметить, что в интервале расстояний  $\mu_0 R = 2,03 + 8,12$  спектры изменяются весьма мало, что свидетельствует об установлении равновесия.

Для оценки точности полученных результатов расчета дополнительно вычислялась величина полной интенсивности излучения, проходящего через сферу радиусом  $R_k$ , а также средняя ошибка этой величины. Исходя из этого, точность вычисления значений факторов накопления можно оценить в 2–5 % для расстояний, равных 250 и 500 м, и в 15 % для расстояния 1000 м. Оценка точности временных угловых и прочих распределений интенсивности производилась сравнением полученных величин при числе испытаний, равном 500, 1000 и 1500. Исходя из такого сравнения, можно сделать вывод, что точность вычисления временного хода интенсивности до  $t = 1$  мкс составляет около 15–20 %, а после  $t = 1$  мкс — около 40–50 %.

Точность вычисления временных энергетических спектров на расстояниях 250 и 500 м от источника до времени  $t = 1 - 1,5$  мкс составляет около 20–30 %; после  $t = 1,5$  мкс спектры следует рассматривать как ориентировочные, так как ошибка расчета может составлять 50 % и более. То же самое можно сказать относительно временного хода интенсивности в различных телесных углах.

<sup>1)</sup> Уже после направления статьи в редакцию была опубликована работа [8], в которой изучается угловое распределение интенсивности рассеянного излучения от источника  $\text{Co}^{60}$ , помещенного в полубесконечную водную среду. Сравнение результатов настоящей работы, относящихся к переднему полупространству, и результатов работы [8] показывает хорошее совпадение значений угловых распределений интенсивности рассеянного излучения. Максимальное расхождение в доле излучения, идущего в данном телесном угле, относительно рассеянного излучения, идущего из всего переднего полупространства составляет  $\sim 25\%$  (для  $\mu_0 R = 4,06$  и интервала углов  $\theta$ , равного  $40-90^\circ$ ), что можно объяснить различием в геометрии опыта [8] (полубесконечное пространство) и расчета (бесконечное пространство). (Прим. при корректуре.)

Точность угловых распределений интенсивности и угловых энергетических спектров, полученных для постоянного источника, составляет около 30%.

В заключение авторы приносят глубокую благодарность И.М. Гельфанду за внимание к работе и содействие в ее проведении.

#### СПИСОК ЛИТЕРАТУРЫ

1. Защита ядерных реакторов/Под ред. Т. Роквелла. — М.: ИЛ, 1958. — Гл. 10.
2. Зигбан К. Бета- и гамма-спектроскопия. — М.: ИЛ, 1959. — Приложение I.
3. Гельфанд И.М. и др. Тр. Второй международной конференции по мирному использованию атомной энергии (Женева, 1958). Докл. советских ученых. Т. 2. — М.: Атомиздат, 1959. — С. 628.
4. Berger M. // J. Res. Nat. Bur. Standards. 1955. V. 55, № 6. P. 343.
5. Коробов Н. М. // Докл. АН СССР. 1957. Т. 115, № 6. С. 1062.
6. Goldstein H., Wilkins J. // Rept. U. S. Atomic Energy Comm. 1955. № 40. P. 3075.
7. Кухтевич В.И., Цыпин С.Г., Шеметенко Б.П. // Атомная энергия. 1958. Т. 5, вып. 6. С. 638.
8. Казанский Ю.А. // Атомная энергия. 1960. Т. 8, вып. 5. С. 432.

## О СПЕКТРОМЕТРЕ БЫСТРЫХ НЕЙТРОНОВ С БОРОВОДОРОДНЫМ СЦИНТИЛЛЯТОРОМ \*)

Известно несколько типов сцинтилляционных спектрометров быстрых нейтронов; среди них наибольшее распространение получили спектрометры с двумя датчиками [1] и спектрометры с одним водородсодержащим сцинтиллятором [2]. В спектрометрах первого типа энергия нейтронов определяется по энергии протонов, рассеянных на определенный угол, так что  $E_n = E_p / \sin^2 \vartheta$ . Спектрометр можно применять для измерения энергии только в коллимированном потоке. Второй тип спектрометров можно применять без коллиматора, но так как нейтронный спектр получается дифференцированием амплитудного распределения, то точность его в ряде измерений не очень велика. Влияние  $\gamma$ -фона в спектрометрах первого типа исключается путем введения в схему спектрометра задержанных совпадений. В спектрометрах второго типа дискриминация  $\gamma$ -фона осуществляется за счет различия во времени высвечивания сцинтиллятора при возбуждении его протонами и электронами.

В работе [3] предложен сцинтилляционный спектрометр с одним датчиком, в состав сцинтиллятора которого введены водород и бор. На основании краткого описания, приведенного в [3], можно представить следующий принцип действия спектрометра. Быстрые нейтроны источника, попадая в сцинтиллятор, после нескольких рассеяний становятся медленными и захватываются ядрами бора. В результате рассеяний и захвата образуются два импульса: первый обусловлен

\*) Приборы и техника эксперимента. 1963. № 4. С. 39–45 (совм. с В.Н. Ававаевым, Ю.А. Егоровым, Ю.В. Орловым, А.С. Фроловым).

протоном отдачи, второй —  $\alpha$ -частицей. Между моментами появления первого и второго импульсов существует некоторое запаздывание. Схема спектрометра строится таким образом, чтобы регистрировались только те нейтроны источника, которые передали сцинтиллятору всю или почти всю свою энергию. Принципиальная блок-схема такого спектрометра показана на рис. 1.

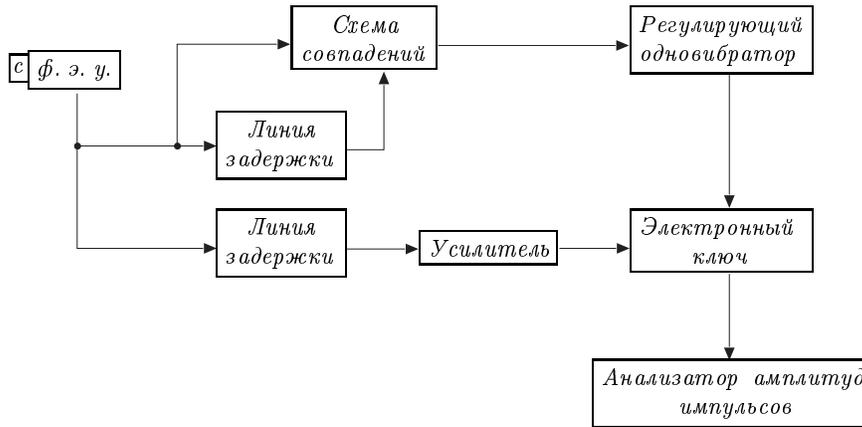


Рис. 1. Принципиальная блок-схема спектрометра быстрых нейтронов с бороводородным сцинтиллятором

Как ясно из рассмотрения принципа действия спектрометра, он может быть использован для измерения нейтронных спектров без коллиматора и при обработке аппаратных спектров не вносятся больших ошибок. Эти достоинства делают спектрометр с бороводородным сцинтиллятором удобным прибором для изучения спектров прохождения нейтронов в “бесконечной” и в “барьерной” геометриях. Для реализации такого спектрометра необходимо определить некоторые его характеристики. Расчет был сделан методом случайных испытаний (методом Монте-Карло) на электронной вычислительной машине.

При расчете характеристик спектрометра рассматривалась следующая задача. Сцинтиллятор представляет собою сферу диаметром  $D$  или цилиндр диаметром и высотой  $D$ . В состав сцинтиллятора входят вещества, состоящие из ядер водорода, углерода и бора. Несмотря на то, что реальный состав сцинтиллятора обычно содержит кислород, взаимодействиями нейтронов с ядрами кислорода пренебрежем. По диаметру сферы (по оси цилиндра) в сцинтиллятор входит поток нейтронов с энергией  $E_{n0}$ . Предполагалось, что нейтрон, попав в сцинтиллятор, на ядрах водорода и углерода только рассеивается, а ядрами бора только поглощается, а также, что, став тепловым, нейтрон не меняет энергии. Сечение поглощения нейтрона ядрами бора во всем диапазоне рассматриваемых энергий изменяется как  $1/\sqrt{E_n}$ .

Задача решена для  $D = 40$  мм и  $D = 80$  мм, трех составов сцинтиллятора (табл. 1) и ряда значений начальной энергии нейтронов в интервале от 0,5 до 10 МэВ. Для удобства энергия нейтронов в расчете представлялась летаргией.

Таблица 1

## Составы сцинтилляторов

№	Вещества в составе	Ядра		
		Н	В	С
I	Ксилол и триметилборат (C <sub>8</sub> H <sub>10</sub> ) + В(ОН <sub>3</sub> ) <sup>*</sup>	$5,7 \cdot 10^{22}$	$3 \cdot 10^{21}$	$3,3 \cdot 10^{22}$
II	Ксилол и триметилборат при обогащении бора изотопом В <sup>10</sup> до 80%	$6,8 \cdot 10^{22}$	$1,26 \cdot 10^{21}$	$2,4 \cdot 10^{22}$
III	Фенилциклогексан и триизобутилборат (C <sub>10</sub> H <sub>16</sub> ) + В(ОС <sub>4</sub> H <sub>9</sub> ) <sub>3</sub>	$5,7 \cdot 10^{22}$	$1,2 \cdot 10^{22}$ <sup>**</sup>	$3,3 \cdot 10^{22}$

<sup>\*</sup>) Смеси равных количеств по объему.  
<sup>\*\*</sup>) Предполагалось, что увеличена концентрация бора.

Были определены следующие величины:

- 1) вероятность поглощения нейтрона ядром бора после рассеяния на ядрах водорода и углерода;
- 2) время от момента первого рассеяния нейтрона на ядре водорода до момента поглощения его ядром бора;
- 3) энергия нейтрона в момент поглощения его ядром бора (только для тех нейтронов, которые сначала рассеивались на ядрах водорода или углерода);
- 4) доля энергии нейтронов, переданная ими в результате рассеяний ядрам водорода (только для нейтронов, которые потом поглотились ядрами бора).

## Методика расчета

Так как рассматриваемый процесс движения нейтронов в сцинтилляторе является процессом без последствия, то траектория их движения конструировалась, как “цепочка” взаимодействий в шестимерном фазовом пространстве  $(X, Y, Z, \omega, E_n)$ , где  $X, Y$  и  $Z$  — координаты,  $\omega$  — направление движения нейтронов между взаимодействиями.

Результаты предварительных расчетов показали, что большинство нейтронов выходит из сцинтиллятора, не взаимодействуя с ядрами водорода или не захватываясь бором [4], поэтому для получения максимальной информации по схеме “условный вылет” моделировалось движение фиктивной частицы с некоторой массой. Движение фиктивной частицы происходит все время внутри объема сцинтиллятора, рассеивается “частица” по законам взаимодействия нейтронов, при каждом акте взаимодействия “частица” теряет часть своей массы [5]. Такая фиктивная частица, вообще говоря, “бессмертна”, но рассмотрение ее прекращалось тогда, когда фиктивная масса становилась меньше некоторой заданной величины  $\varepsilon$ , выбранной из условия обеспечения необходимой точности.

Траектория движения “частицы” в сцинтилляторе конструировалась следующим образом:

- а) рождение “частицы” (состояние  $n = 0$ );
- б) движение “частицы” от столкновения к столкновению (переход из состояния  $n$  в состояние  $n + 1$ ); между столкновениями частица движется равномерно; направление движения  $\omega$  разыграно; рассматривается вероятность выхода “частицы” из сцинтиллятора  $\Omega$  и определяется длина свободного пробега  $S_n$ ;

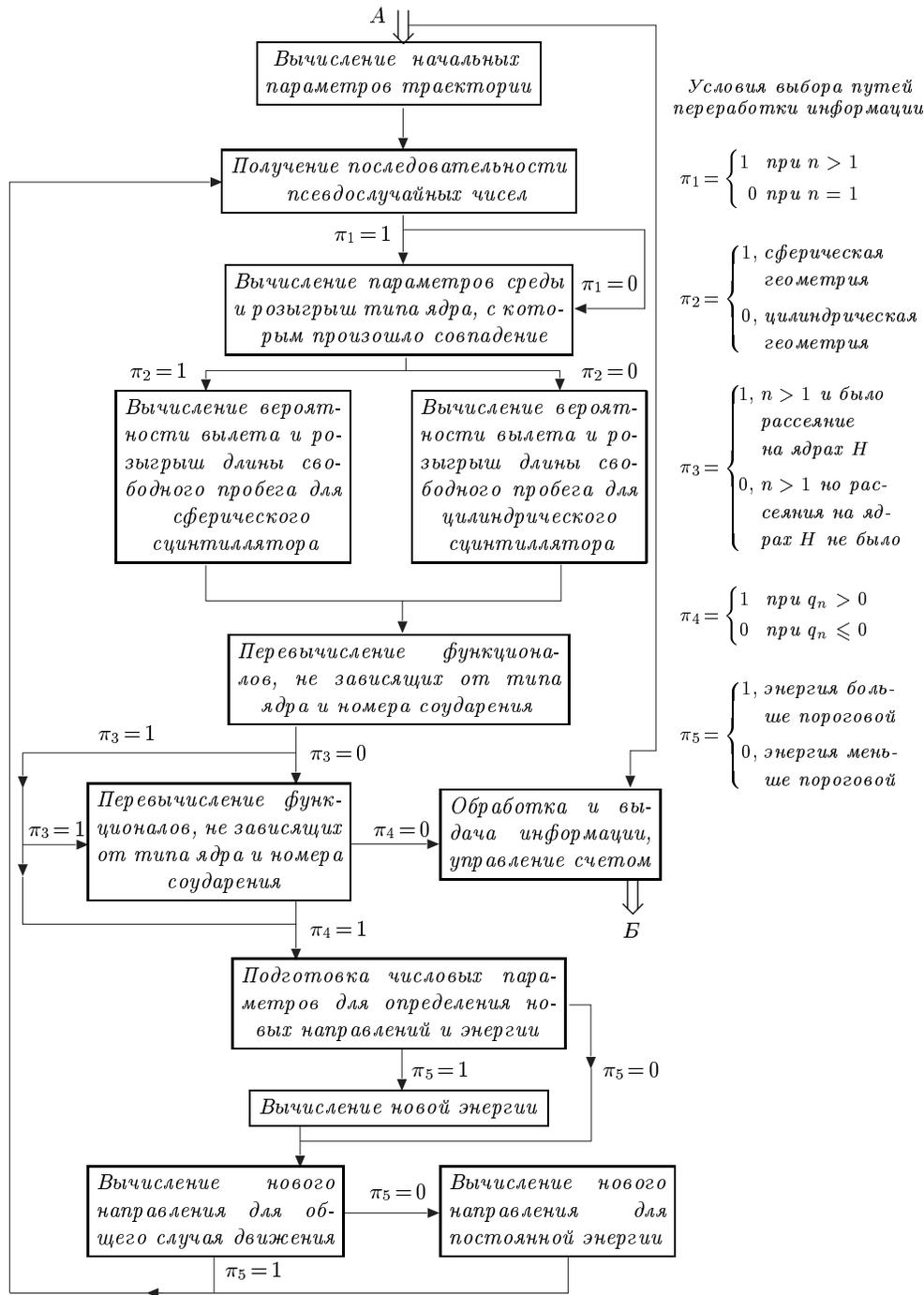


Рис. 2. Блок-схема программы и условия выбора путей переработки информации (A — начало счета, B — конец счета; стрелками показано направление прохождения сигналов)

- в) столкновение “частицы”; находятся координаты  $(n + 1)$ -го столкновения; вероятность рассеяния  $\gamma$  и фиктивная масса после столкновения  $m_{n+1}$ ;
- г) разыгрывается новое направление движения “частицы”;
- д) окончание траектории.

Переход “частицы” из состояния  $n$  в состояние  $n + 1$  рассматривался в такой последовательности. Свободный пробег определялся из известного соотношения  $F(S_n) = \exp(-\Sigma_n S_n)$ , где  $\Sigma_n$  — макроскопическое сечение вещества сцинтиллятора для энергии  $E_n$ . Вероятность выйти частице за пределы сцинтиллятора при движении в направлении  $\omega$   $\Omega_n = \exp(-\Sigma_n L_n)$  ( $L_n$  — расстояние от границы сцинтиллятора до точки  $n$ -го рассеяния). Если “частица” не вышла за пределы сцинтиллятора, то ее свободный пробег распределен по условному закону

$$P(S_n) = \frac{\exp(-\Sigma_n S_n)}{1 - \exp(-\Sigma_n L_n)}.$$

При столкновении фиктивная масса изменяется в соответствии с выражением  $m_{n+1} = m_n \gamma_{n+1} (1 - \Omega_n)$ , где  $\gamma_{n+1}$  — вероятность взаимодействия при  $n$ -м столкновении. При рассмотрении полной траектории “частицы” определяется функционал

$$\varphi = \sum_n^{\infty} (1 - \gamma_{n-1}) m_n (1 - \Omega_n).$$

Суммирование распространяется на все столкновения, причем  $m_0 = 1$ . При помощи  $\varphi$  определяются основные статистики.

Эффективность спектрометра  $U$ , например, рассчитывается как математическое ожидание функционала  $\varphi$  от траектории нейтрона  $\xi(T)$ :  $U = M\varphi[\xi(T)]$ . Распределение нейтронов, которые после рассеяний на ядрах водорода или углерода были захвачены бором по энергетическим группам  $R_n(E_n)$ , определяется как  $R_n(E_n) = \{r_n^{(i)}\}_{i=1,2,\dots,i_{\text{кон}}}$ , где  $i$  — номер энергетического интервала, а  $r_n^{(i)} = M\psi^{(i)}[\xi(T)]$ . Аналогичным образом были рассчитаны и другие величины. Было определено более 12 величин, позволивших во всей полноте судить о физическом процессе в сцинтилляторе.

Блок-схема программы, по которой выполнялся расчет характеристик спектрометра, показана на рис. 2, где также указаны условия выбора путей переработки информации. При расчете использованы псевдослучайные числа типа предложенных Н.М. Коробовым. Вычисления основных величин были сделаны с точностью не хуже 15 %.

### Анализ результатов расчета

При рассмотрении результатов расчета было установлено, что полное число нейтронов, поглощаемых сцинтиллятором как в результате первого взаимодействия, так и после рассеяний, изменяется в зависимости от энергии нейтронов  $E_{n0}$  по закону

$$P = \frac{CV}{E_{n0}},$$

где  $V$  — объем сцинтиллятора;  $C$  — величина, зависящая от состава сцинтиллятора.

Истинная эффективность спектрометра  $U$ , определенная как отношение числа нейтронов, захваченных бором (хотя бы после одного рассеяния на ядре водорода), к потоку нейтронов, падающему на сцинтиллятор, равна

$$U = P - [1 - \exp(-\Sigma_a D)],$$

где  $\Sigma_a$  — макроскопическое сечение поглощения нейтронов ядрами бора в сцинтилляторе,  $D$  — размер сцинтиллятора в направлении потока нейтронов. Полученные закономерности показаны на рис. 3, из которого видно, что значения  $P$

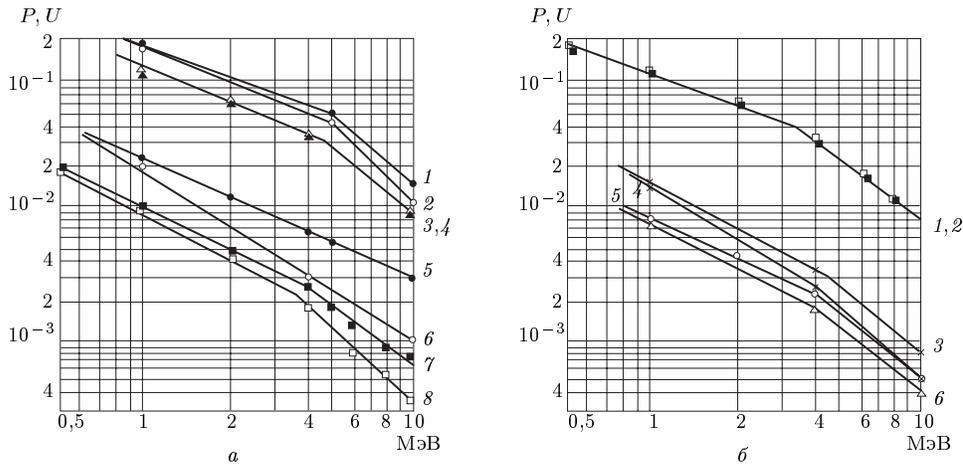


Рис. 3. Зависимость эффективности спектрометра от энергии нейтронов. *a*: 1, 2 — сфера  $\varnothing$  80 мм, II состав, 3, 4 — сфера  $\varnothing$  80 мм, III состав, 5, 6 — сфера  $\varnothing$  40 мм, II состав, 7, 8 — сфера  $\varnothing$  40 мм, I состав; *b*: 1, 2 — сфера  $\varnothing$  80 мм, I состав, 3, 4 — цилиндр  $\varnothing$  40  $\times$  40 мм, I состав, 5, 6 — сфера  $\varnothing$  40 мм, III состав

в зависимости от энергий нейтронов изменяются (кривая 7) примерно так же, как микроскопическое сечение рассеяния нейтронов на водороде. Истинная эффективность  $U$  мало отличается от  $P$ , а расхождение кривых 7 и 8 при больших энергиях обусловлено влиянием захватов ядрами бора. Эффективность  $P$  для состава сцинтиллятора с обогащенным бором не следует зависимости  $\sigma_H(E_{n0})$ , что объясняется высокой концентрацией бора в сцинтилляторе. Для сцинтилляторов больших размеров эффективность определяется в основном многократными рассеяниями нейтронов в объеме сцинтиллятора, а полная эффективность следует сечению рассеяния на водороде.

Результаты расчета позволили выяснить временные соотношения для процессов в сцинтилляторе, при рассмотрении которых принималось, что первое рассеяние на водороде происходит при  $t = 0$ . Было установлено, что при заданных объеме и составе сцинтиллятора временное распределение процессов “рассеяние–захват” изменяется как  $t^{-k}$  и не зависит от энергии нейтронов (рис. 4). Незначительная зависимость  $t = f(E_{n0})$  при  $t < 0,05$  мкс, по-видимому, связана с некоторыми приближениями, принятыми при расчете. При увеличении размеров

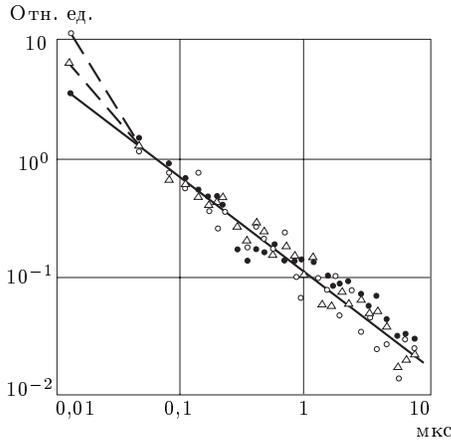


Рис. 4. Временное распределение процессов “рассеяние–захват” в зависимости от времени ( $\Delta U$  — эффективность, определяющая процессы “рассеяние–захват” в  $i$ -м интервале времени). Ось ординат — число регистрации в единичном интервале времени, ось абсцисс — время от начала первого соударения. ● — 1 МэВ,  $\Delta$  — 4 МэВ,  $\circ$  — 10 МэВ (сфера  $\varnothing$  40 мм; I состав)

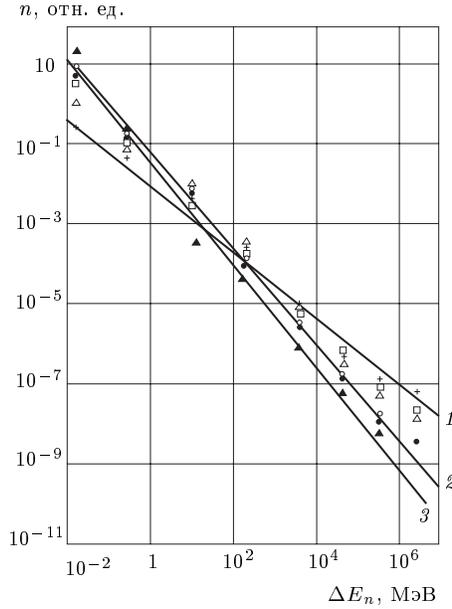


Рис. 5. Энергетическое распределение нейтронов, захваченных бором после рассеяний на водороде; нормированное на интервал энергии 1 эВ ( $E_{n0}$ , МэВ): 1 — II состав, 2 — I состав, 3 — III состав

сцинтиллятора коэффициент  $k$  уменьшается. С увеличением в сцинтиляторе концентрации ядер бора возрастает число процессов “рассеяние–захват” в малые интервалы времени после первого рассеяния.

На рис. 5 показано энергетическое распределение нейтронов в единичном интервале энергий, участвовавших в процессах “рассеяние–захват”. Из рисунка видно, что подавляющее число нейтронов захватывается ядрами бора после того, как они передали почти всю свою энергию ядрам водорода или углерода. Энергетическое разрешение спектрометра, связанное с принципом его действия, весьма высокое.

Полученные закономерности позволяют определить оптимальное время задержки в управляющем канале спектрометра, разрешающее время схемы совпадений и допустимую загрузку спектрометра.

Для получения надежных результатов при большой загрузке спектрометра ( $\sim 10^5 \text{ см}^{-2}\text{с}^{-1}$ ) необходимо выбирать время задержки и разрешающее время схемы совпадений такими, чтобы отношение полезных отсчетов к ложным было максимальным. Из анализа временного распределения следует, что для этого время задержки и разрешающее время схемы совпадений должны быть выбраны возможно минимальными. Однако в случае измерений в малых потоках нейтронов ( $\sim 10^3 \text{ см}^{-2}\text{с}^{-1}$ ) можно выбрать оптимальное разрешающее время схемы совпадений и время задержки из условий обеспечения максимального числа полезных отсчетов  $N_2$  при постоянном отношении их к числу ложных совпадений  $N_1$ , т.е. из условия  $N_1 = \alpha N_2$  ( $\alpha \ll 1$ ). В то же время известно, что  $N_1 = \beta \tau N^2$ ,

где  $\beta$  — вероятность регистрации любого взаимодействия и  $\tau$  — разрешающее время схемы совпадений. С другой стороны,  $N_2 = QN$ , где

$$Q = \sum_s^i U_i, \quad i = 4, 5, 6, \dots$$

Здесь  $U_i$  — число регистраций событий “рассеяние-захват” в  $i$ -м интервале времени. Из записанных соотношений легко получить, что

$$N_2 = \frac{\alpha Q^2}{\beta \tau}. \quad (1)$$

Выбирая  $t_0 = 0,1$  мкс, делая расчет по формуле (1), можно получить зависимости, показанные на рис. 6. Сравнивая эти кривые между собой, можно

Отн. ед.

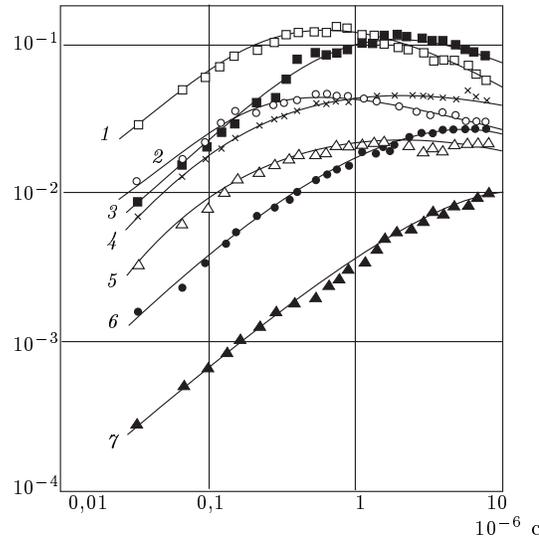


Рис. 6. Эффективность спектрометра в зависимости от разрешающего времени схемы совпадений: 1 — сфера  $\varnothing$  40 мм, II состав; 2 — сфера  $\varnothing$  40 мм, I состав; 3 — сфера  $\varnothing$  80 мм, II состав; 4 — цилиндр  $\varnothing$  40  $\times$  40 мм, I состав; 5 — сфера  $\varnothing$  40 мм, III состав; 6 — сфера  $\varnothing$  80 мм, I состав; 7 — сфера  $\varnothing$  80 мм, III состав

Таблица 2

Условия работы спектрометра с бороводородным сцинтилятором

Состав			I		II		III			
Диаметр кюветы, мм			80		40		80		40	
$\tau$ , мкс			5		1,5		1,5		0,5	
Энергия нейтронов, мэВ	1	$N_1$	$5 \cdot 10^3$	$1 \cdot 10^4$	$1,5 \cdot 10^3$	$3 \cdot 10^4$	$1,5 \cdot 10^4$	$6,7 \cdot 10^3$	$9,4 \cdot 10^3$	
		$N_2$	180	166	3,2	$1,2 \cdot 10^3$	120	67	1,3	
	4	$N_1$	$2 \cdot 10^3$	$3 \cdot 10^3$	$3 \cdot 10^2$	—	—	—	—	
		$N_2$	20	14	0,14	—	—	—	—	
	10	$N_1$	$5 \cdot 10^2$	$8 \cdot 10^2$	$6,4 \cdot 10^1$	$3 \cdot 10^3$	$7,6 \cdot 10^2$	$5,7 \cdot 10^2$	$4,8 \cdot 10^1$	
		$N_2$	1,3	0,96	$0,6 \cdot 10^2$	14	2,9	0,5	$0,3 \cdot 10^2$	

заметить, что эффективность спектрометра в зависимости от разрешающего времени схемы совпадений имеет максимум, причем положение максимума зависит как от размеров, так и от состава сцинтиллятора. Выбирая  $Q$ , соответствующее максимуму, можно рассчитать допустимые потоки монохроматических нейтронов и соответствующие им количества полезных совпадений. Результаты расчета для  $\alpha/\beta = 1$  даны в табл. 2, из которой видно, что спектрометр с бороводородным сцинтиллятором может работать только в относительно небольших нейтронных потоках.

### Заключение

Результаты расчета и анализа характеристик сцинтилляционного спектрометра быстрых нейтронов с бороводородным сцинтиллятором показали, что из всех рассмотренных составов наиболее подходящим для спектрометра является состав II, т.е. смесь равных количеств ксилола (или фенилциклогексана) с триметилборатом при обогащении бора  $B^{10}$  до 80 %, которая залита в кювету  $\varnothing 80$  мм. Разрешающее время схемы совпадений при использовании этого сцинтиллятора должно быть 1,5 мкс.

Результаты расчета позволили сделать вывод, что блок-схема спектрометра

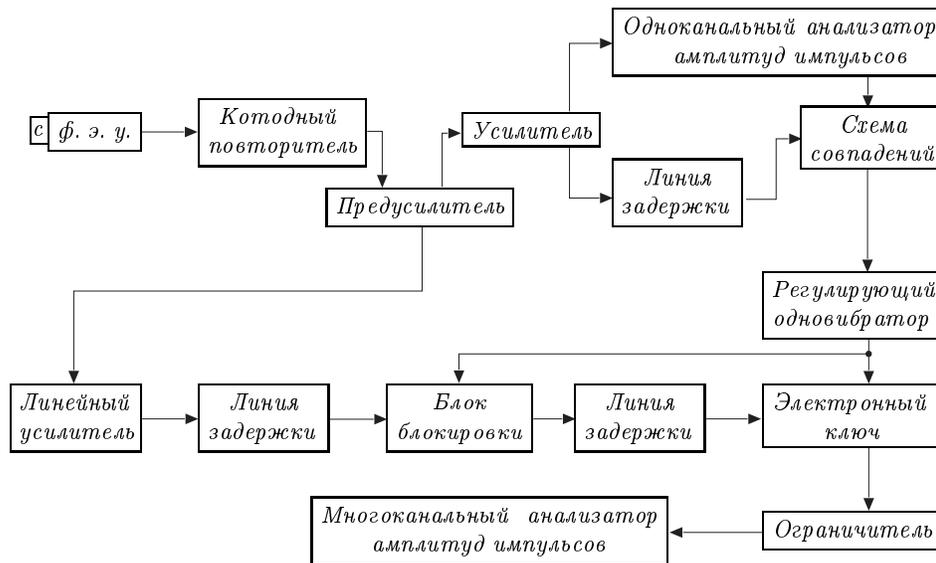


Рис. 7. Предполагаемая блок-схема спектрометра с бороводородным сцинтиллятором

должна быть такой, как показано на рис. 7, причем для снижения числа случайных совпадений должен быть введен одноканальный анализатор амплитуд импульсов в управляющий канал спектрометра и блок блокировки — в главный канал.

## СПИСОК ЛИТЕРАТУРЫ

1. *Beghian L., Allen R., Calvert J., Habban H.* // *Phys. Rev.* 1952. V. 86. P. 1044; *Chagnon R.R., Madancky L., Owen G.E.* // *Rev. Scient. Instrum.* 1953. V. 24. P. 656; *Draper J.E.* // *Rev. Scient. Instrum.* 1954. V. 25. P. 558; *Егоров Ю.А.* // *ПТЭ.* 1958. № 3. С. 21.
2. *Owen R.B.* // *Trans. I. R. E. Nucl. Sci.* 1958. S.-5. 198; *Brooks F.D.* // *Nucl. Inst.* 1959. V. 4. P. 151; *Егоров Ю.А., Панкратьев Ю.В.* // *Инж.-физ. ж.* 1961. Т. 4. С. 112. *Broek H.W.* // *Rev. Scient. Instrum.* 1960. V. 31. P. 1063; *Дулин В.А., Казанский Ю.А., Кузнецов В.Д., Смиренин Г.Н.* // *ПТЭ.* 1961. № 2. С. 35.
3. *Marshall R.S.* // *Bull. Amer. Phys. Soc.* 1952. V. 27. P. 11.
4. *Аваев В.А., Егоров Ю.А., Орлов Ю.В.* // *Вопросы физики защиты реакторов.* — М.: Атомиздат, 1963.
5. *Фролов А.С., Ченцов Н.Н.* // *Тр. VI Всесоюзного совещания по теории вероятностей и матем. статистики, Вильнюс, 1961.* — С. 425–437.

## МОДЕЛИРОВАНИЕ ПРОЦЕССА РАССЕЯНИЯ ЧАСТИЦ В МЕТОДЕ МОНТЕ-КАРЛО<sup>\*)</sup>

Для задач теории переноса необходимой частью алгоритмов Монте-Карло является моделирование направления полета частицы после столкновения. При анизотропном рассеянии координаты нового направления оказываются сложными функциями координат направления до столкновения. Приводимые в руководствах по методу Монте-Карло формулы моделирования содержат особенности [1, 2]. Во избежание “хвостов” в программах расчета приходится переходить к специальным формулам в окрестности особых точек [3]. Причина возникновения особенностей связана с теоремой Пуанкаре–Брауэра о том, что на сфере четной размерности не существует непрерывного касательного векторного поля [4]. Законы анизотропного рассеяния формулируются относительно подвижной системы координат, связанной с направлением полета частицы до столкновения. Построение подвижной системы, т.е. определение матрицы перехода от одной системы координат к другой, требует построения двух полей касательных направлений к единичной сфере в пространстве  $\mathbb{R}^3$ . Попытка решить эту задачу в классе непрерывных функций и приводит к появлению неустранимых особенностей. В работе показано, как отмеченные трудности преодолеваются при построении подвижной системы координат с использованием разрывных функций. Для задач с азимутальной анизотропией (перенос поляризованного излучения) развитая техника не только избавляет от особенностей, но и позволяет получить более экономичные алгоритмы моделирования рассеяния по сравнению с опубликованными ранее.

Пусть  $\vec{\omega}_0 = (\alpha_0, \beta_0, \gamma_0)$  — единичный вектор направления скорости частицы,

<sup>\*)</sup> ВАНТ. Сер. Методики и программы численного решения задач математической физики. 1982. В. 1 (9). С. 3–9 (совм. с А.Н. Субботиным).

$\vec{i}, \vec{j}, \vec{k}$  — орты неподвижной системы координат в пространстве  $\mathbb{R}^3$ ,  $\alpha_0^2 + \beta_0^2 + \gamma_0^2 = 1$ .

Свяжем с  $\vec{\omega}_0$  подвижную систему координат с ортами  $\vec{i}^{\prime}, \vec{j}^{\prime}, \vec{k}^{\prime}$  так, что  $\vec{k}^{\prime} = \vec{\omega}_0$ . Обозначим новое направление полета  $\vec{\omega} = (\alpha, \beta, \gamma)$ ;  $(\vartheta, \varphi)$  — его полярные координаты в подвижной системе,  $\mu = \cos \vartheta = \vec{\omega}_0 \cdot \vec{\omega}$ .

Для неполяризованных частиц рассеяние цилиндрически симметрично относительно старого направления движения  $\vec{\omega}_0$ , т.е. азимутальный угол  $\varphi$  считается распределенным равномерно на  $[0, 2\pi]$ , а для  $\mu = \cos \vartheta$  задается плотность распределения  $g(\mu)$  — индикатриса рассеяния:

$$g(\mu) \geq 0, \quad -1 \leq \mu \leq 1, \quad \int_{-1}^1 g(\mu) d\mu = 1.$$

Так ставятся многие задачи переноса нейтронов и гамма-квантов. В этом случае плотность распределения  $p(\mu, \varphi)$  представляется в виде произведения функций  $\mu$  и  $\varphi$ , что означает независимость случайных величин  $\mu$  и  $\varphi$ :  $p(\mu, \varphi) = g(\mu)/(2\pi)$ ,  $-1 \leq \mu \leq 1$ ,  $0 \leq \varphi \leq 2\pi$ .

В соответствии с этим моделирование рассеяния проводится следующим образом. Сначала из уравнения

$$\int_{-1}^{\mu} g(\mu) d\mu = \xi$$

( $\xi$  — случайное число,  $\xi \in [0, 1]$ ) определяется  $\mu = \cos \vartheta$ . Затем разыгрывается равномерно распределенный азимутальный угол  $\varphi$  (или непосредственно  $\cos \varphi$  и  $\sin \varphi$ ). Вектор, имеющий в подвижной системе координаты  $\sqrt{1 - \mu^2} \cos \varphi$ ,  $\sqrt{1 - \mu^2} \sin \varphi$ ,  $\mu$ , является искомым вектором нового направления. Остается найти его координаты в неподвижной системе координат. Соответствующая матрица перехода образуется координатами ортов  $\vec{i}^{\prime}, \vec{j}^{\prime}, \vec{k}^{\prime}$ . По условию  $\vec{k}^{\prime} = \vec{\omega}_0$ , а  $\vec{i}^{\prime}$  и  $\vec{j}^{\prime}$  обычно выбирают следующим образом:

$$\begin{aligned} \vec{i}^{\prime} &= \frac{[\vec{k}^{\prime} \times \vec{\omega}_0]}{|\vec{k}^{\prime} \times \vec{\omega}_0|} = \left( \frac{-\beta_0}{\sqrt{1 - \gamma_0^2}}; \frac{\alpha_0}{\sqrt{1 - \gamma_0^2}}; 0 \right), \\ \vec{j}^{\prime} &= [\vec{k}^{\prime} \times \vec{i}^{\prime}] = \left( \frac{-\alpha_0 \gamma_0}{\sqrt{1 - \gamma_0^2}}; \frac{-\beta_0 \gamma_0}{\sqrt{1 - \gamma_0^2}}; \frac{1 - \gamma_0^2}{\sqrt{1 - \gamma_0^2}} \right). \end{aligned} \quad (1)$$

Тогда координаты  $\vec{\omega}$  оказываются следующими:

$$\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = A \begin{pmatrix} \sqrt{1 - \mu^2} \cos \varphi \\ \sqrt{1 - \mu^2} \sin \varphi \\ \mu \end{pmatrix}, \quad (2)$$

где

$$A = \begin{pmatrix} \frac{-\beta_0}{\sqrt{1 - \gamma_0^2}} & \frac{-\alpha_0 \gamma_0}{\sqrt{1 - \gamma_0^2}} & \alpha_0 \\ \frac{\alpha_0}{\sqrt{1 - \gamma_0^2}} & \frac{-\beta_0 \gamma_0}{\sqrt{1 - \gamma_0^2}} & \beta_0 \\ 0 & \frac{1 - \gamma_0^2}{\sqrt{1 - \gamma_0^2}} & \gamma_0 \end{pmatrix}, \quad (3)$$

т.е.

$$\begin{aligned}\alpha &= -\beta_0 \sqrt{\frac{1-\mu^2}{1-\gamma_0^2}} \cos \varphi + \alpha_0 \left( \mu - \gamma_0 \sqrt{\frac{1-\mu^2}{1-\gamma_0^2}} \sin \varphi \right), \\ \beta &= \alpha_0 \sqrt{\frac{1-\mu^2}{1-\gamma_0^2}} \cos \varphi + \beta_0 \left( \mu - \gamma_0 \sqrt{\frac{1-\mu^2}{1-\gamma_0^2}} \sin \varphi \right), \\ \gamma &= \gamma_0 \mu + (1-\gamma_0^2) \sqrt{\frac{1-\mu^2}{1-\gamma_0^2}} \sin \varphi.\end{aligned}\quad (4)$$

Эти хорошо известные формулы имеют неустранимую особенность при  $\gamma_0 = \pm 1$ .

Можно попытаться избежать особенностей, более естественным образом учитывая азимутальную изотропию рассеяния и не строя явно подвижную систему координат.

Смодулируем изотропное направление  $\vec{\omega}'$  с координатами  $(\alpha', \beta', \gamma')$ ,  $|\vec{\omega}'| = 1$ . Это делается легко. Достаточно положить  $g(\mu) = 1/2$  и заметить, что никакого пересчета координат из одной системы в другую не требуется. В силу изотропности азимута рассеяния плоскость, образованная векторами  $\vec{\omega}_0$  и  $\vec{\omega}'$ , определяет плоскость рассеяния. Далее,  $c = (\vec{\omega}_0 \cdot \vec{\omega}') = \alpha_0 \alpha' + \beta_0 \beta' + \gamma_0 \gamma'$  — число, равномерно распределенное на  $[-1, +1]$ , поскольку в системе координат, связанной с  $\vec{\omega}_0$ , это есть косинус полярного угла изотропного направления. Можно искать вектор  $\vec{\omega}$  нового направления в виде линейной комбинации векторов  $\vec{\omega}$  и  $\vec{\omega}'$ :

$$\vec{\omega} = p \vec{\omega}_0 + q \vec{\omega}'.$$

Обозначая, как и раньше,  $\vec{\omega} \cdot \vec{\omega}_0 = \cos \vartheta = \mu$ , получаем систему уравнений для определения  $p$  и  $q$

$$\begin{aligned}(\vec{\omega}_0 \cdot \vec{\omega}) &= \mu = p(\vec{\omega}_0 \cdot \vec{\omega}_0) + q(\vec{\omega}_0 \cdot \vec{\omega}') = p + qC, \\ (\vec{\omega} \cdot \vec{\omega}) &= 1 = p^2 + q^2 + 2pqC.\end{aligned}$$

Отсюда

$$p = \mu - C \sqrt{\frac{1-\mu^2}{1-C^2}}, \quad q = \sqrt{\frac{1-\mu^2}{1-C^2}}. \quad (5)$$

Таким образом, для моделирования координат  $\vec{\omega}$  получаются следующие формулы:

$$\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \left( \mu - C \sqrt{\frac{1-\mu^2}{1-C^2}} \right) \begin{pmatrix} \alpha_0 \\ \beta_0 \\ \gamma_0 \end{pmatrix} + \sqrt{\frac{1-\mu^2}{1-C^2}} \begin{pmatrix} \alpha' \\ \beta' \\ \gamma' \end{pmatrix}. \quad (8)$$

Особенности имеются при  $C = \pm 1$ . Если для моделирования  $\mu$  использовать в качестве случайного числа  $C$ , то при некоторых ограничениях на индикатрису рассеяния  $g(\mu)$  можно избавиться от особенностей. Допустим,  $g(\mu)$  непрерывна и  $0 < g(\mu) < \infty$ . Тогда из уравнения для моделирования  $\mu$

$$\int_{-1}^{\mu} g(\mu) d\mu = \frac{1+C}{2}$$

имеем

$$g(\mu) = \frac{1}{2} \frac{dC}{d\mu}, \quad \mu \rightarrow \pm 1 \quad \text{при} \quad C \rightarrow \pm 1.$$

Вычисляя пределы в выражениях (5), находим

$$\begin{aligned} \lim_{C \rightarrow +1} q &= \frac{1}{\sqrt{2g(1)}}, & \lim_{C \rightarrow +1} p &= 1 - \frac{1}{\sqrt{2g(1)}}, \\ \lim_{C \rightarrow -1} q &= \frac{1}{\sqrt{2g(-1)}}, & \lim_{C \rightarrow -1} p &= -1 + \frac{1}{\sqrt{2g(-1)}}. \end{aligned}$$

Легко заметить, что для устранения особенностей в (6) достаточно выполнения указанных ограничений на индикатрису  $g(\mu)$  в некоторой окрестности точек  $\mu = \pm 1$ .

Особенно изящные формулы получаются для моделирования рассеяния нейтрона на неподвижных ядрах в случае изотропного рассеяния в системе центра масс. Пусть  $x$  — угол рассеяния в системе центра масс;  $M$  — масса рассеивающего ядра. Как известно,  $\mu$  и  $x$  связаны следующим соотношением:

$$\mu = \frac{1 + M \cos x}{\sqrt{M^2 + 2M \cos x + 1}}.$$

Используя  $C$  для моделирования  $x$ , имеем в предположении изотропности рассеяния

$$\cos x = C, \quad \mu = \frac{1 + MC}{\sqrt{M^2 + 2MC + 1}}.$$

Вычисляя  $p$  и  $q$ , получим окончательно

$$\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \frac{1}{\sqrt{M^2 + 2MC + 1}} \begin{pmatrix} \alpha_0 \\ \beta_0 \\ \gamma_0 \end{pmatrix} + \frac{M}{\sqrt{M^2 + 2MC + 1}} \begin{pmatrix} \alpha' \\ \beta' \\ \gamma' \end{pmatrix}. \quad (7)$$

Формулы (7) были опубликованы в одной из ранних работ по методу Монте-Карло [5]. Формулы наглядно демонстрируют известный факт, что при  $M \gg 1$  в случае рассеяния, изотропного в системе центра масс, анизотропия в лабораторной системе мала. Особенность, возникающая при  $M = 1$ ,  $x = \pi$  (невозможность рассеяния “назад” на водороде:  $g(\mu) = 0$  при  $\mu < 0$ ), легко ликвидируется, если считать массу ядра водорода  $M > 1,008$ .

К сожалению, в конкретных задачах моделирование  $\mu$  подчас представляет определенные трудности, и избавление от особенностей требует усложнения и удорожания алгоритма моделирования. Поэтому описанная схема не нашла широкого применения<sup>1)</sup>. Более эффективным оказался следующий подход.

<sup>1)</sup>Формулы типа (6) оказываются выгодными, когда важно, чтобы новое направление гладким образом зависело от разыгрываемых углов и старого направления. Так бывает, когда для вычисления средних от функционалов по траекториям используются детерминированные квадратурные формулы.

Орты  $\vec{i}$  и  $\vec{j}$  подвижной системы нужны лишь постольку, поскольку через их координаты в исходном базисе  $\{\vec{i}, \vec{j}, \vec{k}\}$  определяется ортогональная матрица пересчета координат.

Поэтому рассмотрим симметрию пространства относительно биссектрисы  $b^{(+)}$  плоского угла между  $\vec{k}$  и  $\vec{k}'$ . Это будет собственное движение (поворот на  $180^\circ$  вокруг  $b^{(+)}$ ), переводящее  $\vec{k}$  в  $\vec{k}'$ . Следовательно, его матрица  $O^{(+)}$  в базисе  $\{\vec{i}, \vec{j}, \vec{k}\}$  может быть принята за искомую. Заметим, что элементы матрицы  $P$  оператора проектирования на прямую  $s$  (ненормированным) направляющим вектором  $(\xi, \eta, \zeta)$  выражаются через  $\xi, \eta, \zeta$  рационально:

$$P = \frac{1}{\xi^2 + \eta^2 + \zeta^2} \begin{vmatrix} \xi^2 & \xi\eta & \xi\zeta \\ \eta\xi & \eta^2 & \eta\zeta \\ \zeta\xi & \zeta\eta & \zeta^2 \end{vmatrix},$$

а матрица оператора отражения в этой прямой равна  $2P - 1$ . Кроме имеющей направляющий вектор  $\vec{k} + \vec{k}'$  биссектрисы  $b^{(+)}$  внутреннего угла между  $\vec{k}$  и  $\vec{k}'$ , имеется биссектриса  $b^{(-)}$  внешнего плоского угла с направляющим вектором  $\vec{k} - \vec{k}'$ . Отражение вектора  $\vec{k}$  в ней приводит к вектору  $-\vec{k}'$ . Так как  $\vec{k}' = (\alpha_0, \beta_0, \gamma_0)$ , то матрицы указанных отражений суть

$$O^{(\pm)} = \begin{vmatrix} \frac{\alpha_0^2}{1 \pm \gamma_0} - 1 & \frac{\alpha_0\beta_0}{1 \pm \gamma_0} & \pm\alpha_0 \\ \frac{\alpha_0\beta_0}{1 \pm \gamma_0} & \frac{\beta_0^2}{1 \pm \gamma_0} - 1 & \pm\beta_0 \\ \pm\alpha_0 & \pm\beta_0 & \pm\gamma_0 \end{vmatrix}, \quad \det O^{(\pm)} = 1.$$

Используем оба движения и введем матрицу

$$A = \begin{vmatrix} 1 - \frac{\alpha_0^2}{1 + |\gamma_0|} & \frac{-\alpha_0\beta_0\delta}{1 + |\gamma_0|} & \alpha_0 \\ \frac{-\alpha_0\beta_0}{1 + |\gamma_0|} & \delta \left(1 - \frac{\beta_0^2}{1 + |\gamma_0|}\right) & \beta_0 \\ -\alpha_0\delta & -\beta_0 & \gamma_0 \end{vmatrix}, \quad (8)$$

где  $\delta = 1$  при  $\gamma_0 \geq 0$ ;  $\delta = -1$  при  $\gamma_0 < 0$ .

Матрица  $A$  ортогональна, так как отличается лишь знаками двух столбцов от одной из матриц  $O^{(\pm)}$ . Неопределенность в матрице  $A$  отсутствует, но при  $\gamma_0 = 0$  три ее элемента терпят разрыв.

Матрицу  $A$  можно получить просто, если заметить, что вектор

$$\vec{\omega}' = (1 - \alpha_0^2 + |\gamma_0|, -\alpha_0\beta_0, -\alpha_0\gamma_0 - \alpha_0\delta)$$

ортогонален вектору  $\vec{\omega} = \vec{k}'$  в силу тождества  $\alpha_0^2 + \beta_0^2 + \gamma_0^2 \equiv 1$ , и положить затем

$$\vec{k}' = \vec{\omega}_0, \quad \vec{i} = \frac{\vec{\omega}'}{|\vec{\omega}'|}, \quad \vec{j} = [\vec{k}' \times \vec{i}]. \quad (9)$$

При вычислении элементов матрицы (8) использовалась операция деления, которая обязана появиться вследствие теоремы Пуанкаре–Брауэра, если не рассматривать какие-либо более сложные функции с особенностями. Из сравнения матриц (3) и (8) видно, что мы не только избавились от особенностей, но и заменили дорогостоящую операцию вычисления корня на вычисление модуля, сохранив одно деление. В этом смысле полученные формулы являются наилучшими. Однако воспользоваться в полной мере этим преимуществом не удастся из-за необходимости вычисления  $\sin \vartheta = \sqrt{1 - \mu^2}$ . Выпишем окончательный вид формул моделирования:

$$\begin{aligned}\alpha &= \sqrt{1 - \mu^2} \cos \varphi + \alpha_0 \left( \mu - \frac{\alpha_0 \sqrt{1 - \mu^2} \cos \varphi + \beta_0 \delta \sqrt{1 - \mu^2} \sin \varphi}{1 + |\gamma_0|} \right), \\ \beta &= \delta \sqrt{1 - \mu^2} \sin \varphi + \beta_0 \left( \mu - \frac{\alpha_0 \sqrt{1 - \mu^2} \cos \varphi + \beta_0 \delta \sqrt{1 - \mu^2} \sin \varphi}{1 + |\gamma_0|} \right), \\ \gamma &= \gamma_0 \mu - \delta (\alpha_0 \sqrt{1 - \mu^2} \cos \varphi + \beta_0 \delta \sqrt{1 - \mu^2} \sin \varphi).\end{aligned}\quad (10)$$

Поскольку в обычных задачах переноса нейтронов и  $\gamma$ -квантов мы можем выбрать произвольно правостороннюю или левостороннюю систему подвижных координат, можно вместо матрицы (8) использовать следующую:

$$A = \begin{vmatrix} 1 - \frac{\alpha_0^2}{1 + |\gamma_0|} & \frac{-\alpha_0 \beta_0}{1 + |\gamma_0|} & \alpha_0 \\ \frac{-\alpha_0 \beta_0}{1 + |\gamma_0|} & 1 - \frac{\beta_0^2}{1 + |\gamma_0|} & \beta_0 \\ -\alpha_0 \delta & -\beta_0 \delta & \gamma_0 \end{vmatrix}.\quad (11)$$

Это несколько упрощает алгоритм моделирования. Вводя обозначения  $\delta_0 = \mu$ ,  $\delta_1 = \sqrt{1 - \mu^2} \cos \varphi$ ,  $\delta_2 = \sqrt{1 - \mu^2} \sin \varphi$ , можно записать формулы пересчета, соответствующие, например, матрице (11) в следующем простом виде:

$$\begin{aligned}\alpha &= \alpha_0 \left( \delta_0 - \frac{\alpha_0 \delta_1 + \beta_0 \delta_2}{1 + |\gamma_0|} \right) + \delta_1, & \beta &= \beta_0 \left( \delta_0 - \frac{\alpha_0 \delta_1 + \beta_0 \delta_2}{1 + |\gamma_0|} \right) + \delta_2, \\ \gamma &= \gamma_0 \delta_0 - \delta (\alpha_0 \delta_1 + \beta_0 \delta_2).\end{aligned}\quad (12)$$

Подвижную систему координат, определяемую матрицей (8), оказалось удобно применять также для моделирования акта рассеяния поляризованного излучения.

В последние годы появился целый ряд работ [2, 6, 7], в которых предлагается использовать метод Монте-Карло для расчета переноса поляризованного света. Для задания интенсивности и состояния поляризации излучения, распространяющегося в точке  $\vec{r} \in \mathbb{R}^3$  в направлении  $\vec{\omega}_0$ , используется четырехмерный вектор-параметр Стокса, компоненты которого имеют равномерность интенсивности [8]:

$$\vec{I}(\vec{r}, \vec{\omega}_0) = \begin{pmatrix} I(\vec{r}, \vec{\omega}_0) \\ Q(\vec{r}, \vec{\omega}_0) \\ U(\vec{r}, \vec{\omega}_0) \\ V(\vec{r}, \vec{\omega}_0) \end{pmatrix}.$$

Параметры Стокса задаются в некоторой фиксированной системе координат, связанной с направлением  $\vec{\omega}_0$ . При вращении осей этой системы координат в плоскости, перпендикулярной  $\vec{\omega}_0$ , на угол  $\psi$  по часовой стрелке параметры Стокса преобразуются матрицей поворота  $L(\psi)$ , имеющей вид

$$L(\psi) = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\psi & \sin 2\psi & 0 \\ 0 & -\sin 2\psi & \cos 2\psi & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}. \quad (13)$$

Если  $\vec{n}_0$  — единичный вектор, определяющий некоторое заданное направление в среде в точке  $\vec{r}$ , то рассеяние светового пучка с вектор-параметром Стокса  $\vec{I}(\vec{r}, \vec{\omega}_0)$  определяется следующим образом [8, 9]:

$$\vec{I}(\vec{r}, \vec{\omega}) = A(\vec{r}, \vec{\omega}_0, \vec{\omega}) \vec{I}(\vec{r}, \vec{\omega}_0),$$

где  $\vec{\omega}$  — вектор нового направления скорости;

$$A(\vec{r}, \vec{\omega}_0, \vec{\omega}) = L(\varphi_2) \Gamma(\vec{r}, \mu) L(-\varphi_1);$$

$\mu = (\vec{\omega}_0 \cdot \vec{\omega}) = \cos \vartheta$ ,  $\vartheta$  — угол рассеяния;

$$\begin{aligned} \cos \varphi_1 &= \frac{(\vec{n}_0 \cdot \vec{\omega}) - \mu(\vec{n}_0 \cdot \vec{\omega}_0)}{|\vec{\omega}_0 \times \vec{\omega}| |\vec{n}_0 \times \vec{\omega}_0|}, & \sin \varphi_1 &= \frac{(\vec{n}_0 \cdot \vec{\omega}_0 \cdot \vec{\omega})}{|\vec{\omega}_0 \times \vec{\omega}| |\vec{n}_0 \times \vec{\omega}_0|}, \\ \cos \varphi_2 &= \frac{-(\vec{n}_0 \cdot \vec{\omega}_0) + \mu(\vec{n}_0 \cdot \vec{\omega})}{|\vec{\omega}_0 \times \vec{\omega}| |\vec{n}_0 \times \vec{\omega}|}, & \sin \varphi_2 &= \frac{-(\vec{n}_0 \cdot \vec{\omega}_0 \cdot \vec{\omega})}{|\vec{\omega}_0 \times \vec{\omega}| |\vec{n}_0 \times \vec{\omega}|}. \end{aligned} \quad (14)$$

$L(\varphi)$  — матрица поворота, определяемая по формуле (13),  $\Gamma(\vec{r}, \mu)$  — угловая матрица, определяемая моделью среды. Смысл углов  $\vartheta$ ,  $\varphi_1$  и  $\varphi_2$  виден из рис. 1.

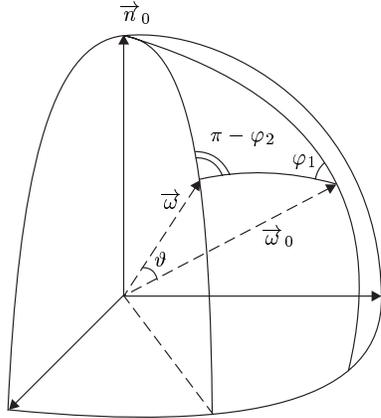


Рис. 1

Матрица  $\Gamma$  определяет закон преобразования  $\vec{I}(\vec{r}, \vec{\omega}_0)$  в акте рассеяния при отнесении значений  $\vec{I}(\vec{r}, \vec{\omega}_0)$  к системе координат, связанной с плоскостью рассеяния, а матрица  $L$  определяет закон преобразования  $\vec{I}(\vec{r}, \vec{\omega}_0)$  при повороте фиксированной системы координат на азимутальный угол, образуемый плоскостью рассеяния в этой фиксированной системе.

При этом  $\vec{\omega}$  является изотропно распределенным. Если  $\vec{\omega}$  выбирается по индикатрисе  $g(\vec{\omega}_0 - \vec{\omega})$  (например, индикатрисе неполяризованного излучения), необходимо

либо перенормировать элементы матрицы рассеяния [2], либо вводить при рассеянии “весовой” множитель  $1/g(\vec{\omega}_0 - \vec{\omega})$ .

Если в качестве  $\vec{n}_0$  выбран единичный вектор, совпадающий по направлению с радиус-вектором  $\vec{r} = (x, y, z)$  точки рассеяния, то в соответствии с формулами (13) и (14) получаются следующие расчетные формулы для моделирования процесса рассеяния [2]:

$$\begin{aligned}
 r = |\vec{r}| &= \sqrt{x^2 + y^2 + z^2}, & \mu_0 &= |\vec{\omega}_0 \times \vec{\omega}| = \sqrt{1 - \mu^2}, \\
 \mu_1 = (\vec{n}_0 \cdot \vec{\omega}_0) &= \frac{1}{r}(\alpha_0 x + \beta_0 y + \gamma_0 z), & \mu_2 &= |\vec{n}_0 \times \vec{\omega}_0| = \sqrt{1 - \mu_1^2}, \\
 \mu_3 = (\vec{n}_0 \cdot \vec{\omega}) &= \frac{1}{r}(\alpha x + \beta y + \gamma z), & \mu_4 &= |\vec{n}_0 \times \vec{\omega}| = \sqrt{1 - \mu_3^2}, \\
 q = (\vec{n}_0 \cdot \vec{\omega}_0 \cdot \vec{\omega}) &= \alpha_0 \beta z + \beta_0 \gamma x + \gamma_0 \alpha y - \alpha_0 \gamma y - \beta_0 \alpha z - \gamma_0 \beta x, \\
 \cos \varphi_1 &= \frac{\mu_3 - \mu \mu_1}{\mu_0 \mu_2}, & \sin \varphi_1 &= \frac{q}{\mu_0 \mu_2}, \\
 \cos \varphi_2 &= \frac{\mu \mu_3 - \mu_2}{\mu_0 \mu_4}, & \sin \varphi_2 &= \frac{-q}{\mu_0 \mu_4}.
 \end{aligned} \tag{15}$$

Сюда еще надо добавить формулы для косинусов и синусов двойных углов, формулы для моделирования  $\vec{\omega}$  типа (4), (10) или (12) и пересчет компонент вектора Стокса после рассеяния

$$\vec{I}(\vec{r}, \vec{\omega}) = L(\varphi_2) \Gamma(\vec{r}, \mu) L(-\varphi_1) \vec{I}(\vec{r}, \vec{\omega}_0).$$

Видно, что основной объем вычислений связан с расчетом элементов матриц поворота  $L(-\varphi_1)$  и  $L(\varphi_2)$ . При этом, даже если пользоваться для моделирования  $\vec{\omega}$  формулами типа (10) или (12), появляются особенности при  $\mu_0 = 0$ ,  $\mu_2 = 0$ ,  $\mu_4 = 0$ . Это связано с тем, что в качестве плоскости отсчета параметров Стокса используется плоскость, образуемая направлением полета и некоторым фиксированным в среде направлением  $\vec{n}_0$ , т.е. система подвижных координат для задания параметров Стокса на траектории светового пучка строится так же, как строился базис  $\vec{i}, \vec{j}, \vec{k}$  (1).

Покажем, что если для задания вектор-параметра Стокса использовать базис (9), то не только исчезают все особенности, но и расчетные формулы становятся заметно экономичнее.

Итак, будем считать, что вектор-параметр Стокса  $\vec{I}(\vec{r}, \vec{\omega}_0)$  для светового пучка, движущегося в направлении  $\vec{\omega}_0$ , вычисляется в системе координат

$$\begin{aligned}
 \vec{i} &= \left( 1 - \frac{\alpha_0^2}{1 + |\gamma_0|}; \frac{-\alpha_0 \beta_0}{1 + |\gamma_0|}; -\alpha_0 \delta \right), \\
 \vec{j} &= \left( \frac{-\alpha_0 \beta_0 \delta}{1 + |\gamma_0|}; \delta \left( 1 - \frac{\beta_0^2}{1 + |\gamma_0|} \right); -\beta_0 \right), \\
 \vec{k} &= (\alpha_0; \beta_0; \gamma_0),
 \end{aligned}$$

где

$$\delta = \begin{cases} 1 & \text{при } \gamma \geq 0, \\ -1 & \text{при } \gamma < 0, \end{cases}$$

и пусть моделирование нового направления  $\vec{\omega} = (\alpha, \beta, \gamma)$  производится по плотности

$$p(\mu, \varphi_1) = \frac{1}{2\pi} g(\mu),$$

где  $g(\mu)$  — индикатриса рассеяния.

Моделируя  $\mu$  и  $\varphi_1$  (или непосредственно  $\cos \varphi_1$  и  $\sin \varphi_1$ ) в системе координат  $\vec{i}^1, \vec{j}^1, \vec{k}^1$ , получаем новое направление  $\vec{\omega} = (\alpha, \beta, \gamma)$  по формулам (10). Нового здесь ничего нет, но следует обратить внимание, что мы пользуемся правосторонней системой координат и что, моделируя  $\cos \varphi_1$  и  $\sin \varphi_1$  и вычисляя  $\cos 2\varphi_1$  и  $\sin 2\varphi_1$ , мы получаем элементы матрицы  $L(-\varphi_1)$ , поскольку  $\varphi_1$  равен углу поворота против часовой стрелки базиса отсчета параметров Стокса  $\vec{i}^1, \vec{j}^1, \vec{k}^1$  к системе координат, связанной с плоскостью рассеяния<sup>2)</sup>. Система координат, в которую должен быть пересчитан вектор Стокса после рассеяния, образуется векторами

$$\begin{aligned} \vec{i}^2 &= \left( 1 - \frac{\alpha^2}{1 + |\gamma|}; \frac{-\alpha\beta}{1 + |\gamma|}; -\alpha\delta \right), \\ \vec{j}^2 &= \left( \frac{-\alpha\beta\delta}{1 + |\gamma|}; \delta \left( 1 - \frac{\beta^2}{1 + |\gamma|} \right); -\beta \right), \\ \vec{k}^2 &= (\alpha; \beta; \gamma). \end{aligned}$$

Чтобы вычислить  $\cos \varphi_2$  и  $\sin \varphi_2$ , обратимся к формальному смыслу матриц

$$L(\varphi_2), \quad \Gamma(\vec{r}^2, \mu), \quad L(-\varphi_1).$$

Матрица  $L(-\varphi_1)$  отвечает преобразованию вектора Стокса при переходе из системы  $\vec{i}^1, \vec{j}^1, \vec{k}^1$  в систему  $\vec{l}_1, \vec{l}_2, \vec{\omega}_0$ , где

$$\vec{l}_2 = \frac{[\vec{\omega}_0 \times \vec{\omega}]}{|\vec{\omega}_0 \times \vec{\omega}|}, \quad \vec{l}_1 = [\vec{\omega}_0 \times \vec{l}_2]. \quad (16)$$

Матрица  $\Gamma(\vec{r}^2, \mu)$  отвечает преобразованию вектора Стокса при переходе из системы  $\vec{l}_1, \vec{l}_2, \vec{\omega}_0$  в систему  $\vec{l}'_1, \vec{l}'_2, \vec{\omega}$ , где

$$\vec{l}'_1 = [\vec{\omega} \times \vec{l}_2]. \quad (17)$$

Матрица  $L(\varphi_2)$  отвечает преобразованию вектора Стокса из системы  $\vec{l}'_1, \vec{l}'_2, \vec{\omega}$  в систему  $\vec{i}^2, \vec{j}^2, \vec{k}^2$ .

Нужно вычислить  $\cos \varphi_2 = (\vec{i}^2 \cdot \vec{l}'_1)$  и  $\sin \varphi_2 = (\vec{j}^2 \cdot \vec{l}'_1)$ , но пока неизвестен вектор  $\vec{l}'_1$ .

Если для вычисления координат вектора  $\vec{l}'_1$  воспользоваться выражения-

<sup>2)</sup> Подобным же обстоятельством пользовались Каттавар и Пласс в работе [7]. Они в качестве  $\vec{n}_0$  выбирали орт  $\vec{k}$  и, по-видимому, пользовались стандартным алгоритмом (формулы (1)–(4)) для моделирования  $\vec{\omega}$ . Это позволило им, как и в нашем случае, сразу вычислить матрицу  $L(-\varphi_1)$ .

ми (16) и (17), то неизбежно появится особенность при  $|\vec{\omega}_0 \times \vec{\omega}| = 0$ . Но легко заметить (рис. 2), что при моделировании  $\cos \varphi_1$  и  $\sin \varphi_1$  строится ненулевой вектор  $\vec{l}'_1$ , координаты которого в системе  $\vec{i}', \vec{j}', \vec{k}'$  равны  $(\cos \varphi_1, \sin \varphi_1, 0)$ . Далее, вектор  $\vec{l}'_1$  лежит в плоскости, образованной векторами  $\vec{\omega}_0$  и  $\vec{\omega}$ ,  $(\vec{\omega} \cdot \vec{l}'_1) = \cos(\vartheta + \pi/2) = -\sin \vartheta = -\sqrt{1 - \mu^2}$  ( $0 \leq \vartheta \leq \pi$ ). Проекция вектора  $\vec{l}'_1$  на плоскость, образуемую векторами  $\vec{i}'$  и  $\vec{j}'$ , при  $\vartheta \leq \pi/2$  имеет направляющие косинусы  $\cos \varphi_1, \sin \varphi_1$ , при  $\vartheta > \pi/2$  соответственно  $-\cos \varphi_1, -\sin \varphi_1$ . Таким образом, координаты вектора  $\vec{l}'_1$  в системе координат  $\vec{i}', \vec{j}', \vec{k}'$  равны

$$\begin{aligned} &(\mu \cos \varphi_1; \mu \sin \varphi_1; -\sqrt{1 - \mu^2}) \quad \text{при } \vartheta \leq \frac{\pi}{2}, \\ &(-\mu(-\cos \varphi_1); -\mu(-\sin \varphi_1); -\sqrt{1 - \mu^2}) \quad \text{при } \vartheta > \frac{\pi}{2}, \end{aligned}$$

т.е. всегда  $(\mu \cos \varphi_1; \mu \sin \varphi_1; -\sqrt{1 - \mu^2})$ . Отсюда получим следующие выражения для координат вектора  $\vec{l}'_1$ :

$$\begin{aligned} &\left( \mu \cos \varphi_1 \left( 1 - \frac{\alpha_0^2}{1 + |\gamma_0|} \right) - \mu \sin \varphi_1 \frac{\alpha_0 \beta_0 \delta}{1 + |\gamma_0|} - \alpha_0 \sqrt{1 - \mu^2}; \right. \\ &\quad \mu \cos \varphi_1 \frac{\alpha_0 \beta_0}{1 + |\gamma_0|} + \mu \sin \varphi_1 \delta \left( 1 - \frac{\beta_0^2}{1 + |\gamma_0|} \right) + \beta_0 \sqrt{1 - \mu^2}; \\ &\quad \left. - \mu \cos \varphi_1 \alpha_0 \delta - \mu \sin \varphi_1 \beta_0 + \gamma_0 \sqrt{1 - \mu^2} \right). \quad (18) \end{aligned}$$

Осталось вычислить

$$\cos \varphi_2 = (\vec{i}'' \cdot \vec{l}'_1), \quad \sin \varphi_2 = (\vec{j}'' \cdot \vec{l}'_1). \quad (19)$$

От особенностей мы избавились. Вместо формул (15) необходимо произвести вычисления по формулам (18) и (19). Организуем программу расчета так, чтобы хранить на траектории координаты векторов  $\vec{i}', \vec{j}', \vec{k}'$  и пользоваться ими не только для моделирования нового направления  $\vec{\omega} = (\alpha, \beta, \gamma)$ , но и для вычисления координат вектора  $\vec{l}'_1$  в точке нового соударения. Тогда объем вычислений оказывается очень небольшим, поскольку в формулах (16) все величины оказываются готовыми.

Конечно, после окончания моделирования всех траекторий вычисленные функционалы придется пересчитать в ту систему отсчета, которая отвечает физическому смыслу измерений. Но это разовая операция для расчета задачи. Точно так же необходимо пересчитать параметры Стокса для моделирования источника. Но обычно источник предполагается естественным, а параметры Стокса такого источника инвариантны относительно поворота системы координат.

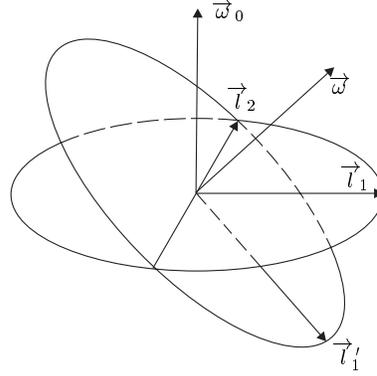


Рис. 2

Рис. 2. Вектор  $\vec{l}'_1$  в системе координат  $\vec{i}', \vec{j}', \vec{k}'$  равен

После того как изложенный в статье геометрический подход был предложен вторым из авторов, использование формул (8) для моделирования неполяризованного излучения было начато И.Ф. Подливаевым и А.С. Фроловым [10]. В дальнейшем эти формулы нашли широкое применение [11, 12]. Распространение этого подхода на моделирование поляризованного излучения принадлежит первому из авторов. Мы благодарны М.Г. Кузьминой, И.Ф. Подливаеву и А.С. Фролову за полезные обсуждения и советы.

#### СПИСОК ЛИТЕРАТУРЫ

1. *Спанье Дж., Гелбард Э.* Метод Монте-Карло и задачи переноса нейтронов. — М.: Атомиздат, 1972.
2. *Марчук Г.И., Михайлов Г.А. и др.* Метод Монте-Карло в атмосферной оптике. — Новосибирск: Наука, 1976.
3. *Бочарова И.Е., Прохорова Л.И., Смиренкин Г.Н.* Расчет методом Монте-Карло детектора нейтронов // Вопросы атомной науки и техники. Сер. Ядерные константы. 1974. Вып. 17.
4. *Александров П.С.* Комбинаторная топология. — М.: Гостехиздат, 1947.
5. *Dismuke N.M.* Monte-Carlo computations // Symposium on Monte-Carlo methods. — N.Y.: Wiley, 1956.
6. *Kattawar G.W., Plass G.N.* Radiance and polarisation of multiple scattered light from haze and clouds // Applied Optics. 1968. V. 7, № 8.
7. *Carter L.L., Horak H.G., Sandford M.T.* An adjoint Monte-Carlo treatment of the equations of radiative transfer for polarized light // Journal of Computational physics. 1978. № 26.
8. *Чандрасекар С.* Перенос лучистой энергии. — М.: ИЛ, 1953.
9. *Кузьмина М.Г.* Уравнение переноса поляризованного излучения в анизотропных средах. — М., 1976. — (Препр. / ИПМ АН СССР; № 68).
10. *Лейпунский О.И., Стрелков А.С., Фролов А.С., Ченцов Н.Н.* Распространение в воздухе  $\gamma$ -излучения из мгновенного точечного источника // АЭ. 1961. Т. 10, вып. 5. С. 483–500.
11. *Марчук Г.И., Михайлов Г.А.* О решении задач атмосферной оптики методом Монте-Карло // Изв. АН СССР. Сер. Физика атмосферы и океана. 1967. Т. 3, № 4.
12. *Марчук Г.И., Михайлов Г.А., Назаралиев М.А., Дарбинян Р.А.* Решение прямых и некоторых обратных задач атмосферной оптики методом Монте-Карло. — Новосибирск: Наука, 1968.

РАЗДЕЛ III

МАТЕМАТИЧЕСКАЯ  
СТАТИСТИКА

---

ОБ АСИМПТОТИЧЕСКОЙ  
ЭФФЕКТИВНОСТИ ОЦЕНКИ  
НАИБОЛЬШЕГО ПРАВДОПОДОБИЯ<sup>\*)</sup>

*Замечание к докладу А.М. Кагана  
“Некоторые свойства оценок максимального правдоподобия”*

В своем интересном сообщении [1] А.М. Каган отметил, что в литературе до сих пор не выяснен вопрос о существовании асимптотически эффективных оценок параметра. Хочу указать, что в моей дипломной работе “Асимптотическая теория статистических оценок” [2], выполненной в 1952 г. под руководством профессора Е.Б. Дынкина, мне удалось доказать, что при выполнении некоторых условий регулярности оценка наибольшего правдоподобия является асимптотически наилучшей в классе асимптотически несмещенных оценок. Это по существу дает ответ на указанный вопрос. Весьма краткое изложение работы было сделано на III Математическом съезде [3]. Ввиду того, что указанный результат имеет некоторый интерес для теории оценок, мне представляется полезным дать здесь его точную формулировку, а также план доказательства.

Рассмотрим последовательность независимых случайных величин с плотностью распределения  $f(x; \alpha)$ , зависящей от некоторого параметра  $\alpha$ . Мы будем предполагать, что параметр  $\alpha$  принимает значения из некоторого отрезка  $[\alpha_1, \alpha_2]$  числовой прямой. При некотором неизвестном значении параметра  $\alpha = \alpha_0$  делаются наблюдения  $x_1, x_2, \dots, x_n, \dots$ . Наша задача — разыскать последовательности функций  $\alpha_1^*(x_1); \dots; \alpha_n^*(x_1, \dots, x_n); \dots$  с минимальным рассеянием от истинного значения параметра  $\alpha_0$ .

Понятие “минимального рассеяния” нуждается в уточнении. Обратимся сначала к случаю, когда оценка  $\alpha^*$  является несмещенной, т.е. когда при любых  $\alpha_0$  и  $n$  выполнено равенство  $M^{\alpha_0} \alpha_n^* = \alpha_0^1$ . Тогда в качестве меры рассеяния около истинного значения параметра обычно принимают дисперсию оценки  $D^{\alpha} \alpha_n^*$  и под оценкой с наилучшими асимптотическими свойствами понимают оценку,

<sup>\*)</sup> Труды VI Всесоюзного совещания по теории вероятн. и матем. статистике (Вильнюс, 1960). — Вильнюс: ГИПНА, 1962. — С. 399–402.

дисперсия  $\mathbf{D}^\alpha \alpha_n^*$  которой имеет как функция  $n$  по возможности наиболее быстрый порядок убывания. Если предположить, лишь незначительно ограничивая общность рассуждений, что оценка  $\alpha^*$  является регулярной в смысле Крамера (см. [4, с. 520]), то из известного неравенства Крамера–Рао будет следовать, что<sup>1)</sup>

$$\mathbf{D}^\alpha \alpha_n^* \geq \frac{1}{nk^2(\alpha_0)},$$

где

$$k^2(\alpha_0) = \mathbf{M}^{\alpha_0} \left[ \frac{\partial}{\partial \alpha} \ln f(x; \alpha) \Big|_{\alpha=\alpha_0} \right]^2.$$

Таким образом, оценки, для которых при всех  $\alpha_0 \in [\alpha_1, \alpha_2]$  имеет место равенство  $\mathbf{D}^{\alpha_0} \alpha_n^* = 1/(nk^2\alpha_0)$ , являются наилучшими в указанном выше смысле. Р. Фишер предложил называть такие оценки эффективными. Позднее было введено понятие асимптотически эффективной оценки, у которой  $\mathbf{D}^{\alpha_0} \alpha_n^* = 1/(nk^2\alpha_0) + o(1/n)$ . Однако до сих пор вопрос о достаточно широких условиях, при которых существуют асимптотически эффективные оценки параметра, остается открытым. Следует отметить, что иногда под асимптотически эффективной оценкой понимают оценку, распределенную асимптотически нормально с параметрами  $(\alpha_0, 1/(k(\alpha_0)\sqrt{n}))$  хотя бы эта оценка не имела ни дисперсии, ни даже математического ожидания. Такое определение связано с другой мерой рассеяния, именно с дисперсией асимптотически эквивалентного нормального распределения, и, как легко видеть, из асимптотической эффективности оценки во втором понимании, не следует ее асимптотическая эффективность в первом понимании.

В общем случае ни эффективных, ни даже несмещенных оценок параметра распределения может не существовать (см., например, [5]). Поэтому мы предлагаем расширить класс допустимых оценок и показываем, что при выполнении некоторых условий типа регулярности в этом классе оценок существуют асимптотически наилучшие.

Пусть оценка  $\alpha^*$  задана последовательностью функций  $\{\alpha_n^*(x_1, \dots, x_n)\}_{n=1}^\infty$ . Будем предполагать, что  $\mathbf{M}^{\alpha_0}(\alpha_n^* - \alpha_0)^2 < \infty$  при всех  $\alpha_0 \in [\alpha_1, \alpha_2]$ .

**ОПРЕДЕЛЕНИЕ.** Назовем оценку  $\alpha_n^*$  *асимптотически несмещенной*, если при любом  $\alpha_0 \in [\alpha_1, \alpha_2]$  имеют место равенства:

- 1)  $\mathbf{M}^{\alpha_0} \alpha_n^* = \alpha_0 + o\left(\frac{1}{\sqrt{n}}\right)$ ;
- 2)  $\frac{d}{d\alpha_0} \mathbf{M}^{\alpha_0} \alpha_n^* = 1 + o(1)$ ;

и в дальнейшем ограничимся рассмотрением только асимптотически несмещенных оценок.

Мы скажем, что оценка  $\alpha^*$  *асимптотически не хуже* оценки  $\beta^*$  и *асимптотически эквивалентна* оценке  $\gamma^*$ , если при любом  $\alpha_0 \in [\alpha_1, \alpha_2]$  соответственно выполнены условия

$$\lim_{n \rightarrow \infty} \frac{\mathbf{M}^{\alpha_0}(\alpha_n^* - \alpha_0)^2}{\mathbf{M}^{\alpha_0}(\beta_n^* - \alpha_0)^2} \leq 1, \quad \lim_{n \rightarrow \infty} \frac{\mathbf{M}^{\alpha_0}(\alpha_n^* - \alpha_0)^2}{\mathbf{M}^{\alpha_0}(\gamma_n^* - \alpha_0)^2} = 1.$$

<sup>1)</sup>  $\mathbf{M}^\alpha \alpha_n^* = \int \alpha_n^*(x_1, \dots, x_n) \prod_{i=1}^n [f(x_i; \alpha) dx_i]$ .

ОПРЕДЕЛЕНИЕ. Мы назовем оценку  $\alpha^*$  *асимптотически наилучшей*, если она асимптотически не хуже любой другой оценки.

В дальнейшем будут использованы нижеследующие условия регулярности.

А.  $f(\cdot, \alpha') \neq f(\cdot, \alpha'')$  при  $\alpha' \neq \alpha''$ .

Б.  $\partial^m \ln f(x; \alpha) / \partial \alpha^m$  существует и  $|\partial^m \ln f(x; \alpha) / \partial \alpha^m| \leq G_m(x)$ , причем  $M^\alpha [G_m(x)]^4 \leq G_m < \infty$ .

В.  $\partial^k f(x; \alpha) / \partial \alpha^k$  существует и  $|\partial^k f(x; \alpha) / \partial \alpha^k| \leq H_k(x)$ , причем  $M^\alpha H_k(x) \leq H_k < \infty$ .

Г.  $M^\alpha [\partial \ln f(x; \alpha) / \partial \alpha]^2 = k^2(\alpha) \geq k^2 > 0$ .

Несколько уточняя обычное определение, будем называть *оценкой наибольшего правдоподобия* оценку  $\varphi^*$ , удовлетворяющую условиям:

1)  $\varphi_n^* = \varphi$ , если

$$\prod_{i=1}^n f(x_i; \varphi) > \prod_{i=1}^n f(x_i; \alpha)$$

при всех  $\alpha \neq \varphi$ ;

2)  $\varphi_n^* = \alpha_1$ , если не существует  $\varphi$ , о котором говорится в 1).

Тогда, уточняя известные результаты (см. [6–8]) и проводя некоторые другие оценки, можно показать следующее.

ЛЕММА 1. Если выполнены условия А, Б<sub>0</sub>, Б<sub>1</sub>, то для всякого  $\varepsilon > 0$  с вероятностью, не меньшей  $1 - c_1(\varepsilon)/n^2$ , максимум функции правдоподобия  $\Pi = \prod_{i=1}^n f(x_i; \alpha)$  лежит в интервале  $(\alpha_0 - \varepsilon, \alpha_0 + \varepsilon)$ .

ЛЕММА 2. Если выполнены условия Б<sub>0</sub>–Б<sub>3</sub>, В<sub>1</sub>, В<sub>2</sub>, то существует такое  $\varepsilon_0 > 0$ , что при  $0 < \varepsilon < \varepsilon_0$  с вероятностью, не меньшей  $1 - c_2(\varepsilon)/n^2$ , в интервале  $(\alpha_0 - \varepsilon, \alpha_0 + \varepsilon)$  существует не более одного корня  $\varphi^*$  уравнения правдоподобия

$$\frac{\partial}{\partial \alpha} \left( \frac{1}{n} \ln \Pi \right) = \frac{1}{n} \sum_{i=1}^n \frac{\partial}{\partial \alpha} \ln f(x_i; \alpha) = 0.$$

ЛЕММА 3. Если выполнены условия Б<sub>0</sub>–Б<sub>5</sub>, В<sub>1</sub>, В<sub>2</sub>, Г, то существует такое  $\varepsilon_0 > 0$ , что при  $0 < \varepsilon < \varepsilon_0$  с вероятностью, не меньшей  $1 - c_3(\varepsilon)/n^2$ , в интервале  $(\alpha_0 - \varepsilon, \alpha_0 + \varepsilon)$  каждый из многочленов

$$\begin{aligned} & \left[ \frac{\partial}{\partial \alpha} \left( \frac{1}{n} \ln \Pi \right) \right]_{\alpha=\alpha_0} + \left[ \frac{\partial^2}{\partial \alpha^2} \left( \frac{1}{n} \ln \Pi \right) \right]_{\alpha=\alpha_0} (\alpha - \alpha_0) + \\ & + \frac{1}{2} \left[ \frac{\partial^3}{\partial \alpha^3} \left( \frac{1}{n} \ln \Pi \right) \right]_{\alpha=\alpha_0} (\alpha - \alpha_0)^2 + \frac{1}{6} \left[ \frac{\partial^4}{\partial \alpha^4} \left( \frac{1}{n} \ln \Pi \right) \right]_{\alpha=\alpha_0} (\alpha - \alpha_0)^3 \pm \\ & \pm \frac{1}{24} \left[ \frac{1}{n} \sum_{i=1}^n G_5(x_i) \right] (\alpha - \alpha_0)^4 \end{aligned}$$

обращается в нуль только один раз соответственно при  $\alpha = \beta$  и  $\alpha = \gamma$ , причем либо  $\beta \leq \varphi^* \leq \gamma$ , либо  $\gamma \leq \varphi^* \leq \beta$ .

ТЕОРЕМА 1. Если выполнены уравнения А, Б<sub>0</sub>–Б<sub>5</sub>, В<sub>1</sub>, В<sub>2</sub>, Г, то оценка наибольшего правдоподобия существует и является регулярной в смысле Крамера, причем при любом  $\alpha_0 \in [\alpha_1, \alpha_2]$ :

- 1)  $\mathbf{M}^{\alpha_0} \varphi_n^* = \alpha_0 + O\left(\frac{1}{n}\right)$ ;
- 2)  $\mathbf{M}^{\alpha_0} (\varphi_n^* - \alpha_0)^2 = \frac{1}{n k^2(\alpha_0)} + O\left(\frac{1}{n\sqrt{n}}\right)$ ;
- 3)  $\frac{d}{d\alpha_0} \mathbf{M}^{\alpha_0} \varphi_n^* = 1 + O\left(\frac{1}{n}\right)$ .

Доказательство. Зададим  $\varepsilon > 0$ , удовлетворяющее условиям лемм 2 и 3. Доопределим величины  $\beta$  и  $\gamma$  посредством равенств  $\beta = \alpha_1$ ,  $\gamma = \alpha_1$  на случай, когда хоть одно из вспомогательных уравнений не имеет единственного корня в  $(\alpha_0 - \varepsilon, \alpha_0 + \varepsilon)$ . Тогда легко вычислить первые два момента величин  $\beta$  и  $\gamma$ , рассматривая их как неявные функции определяющих их уравнений (см. [4, с. 388]). С помощью леммы 3 мы можем показать, что  $\mathbf{M}|\beta - \varphi^*| = O(1/(n\sqrt{n}))$ . Отсюда, принимая во внимание леммы 1 и 2, мы получаем первое и второе утверждение теоремы. Величина  $\frac{d}{d\alpha} \mathbf{M}^{\alpha} \varphi_n^*$  оценивается аналогичным образом при помощи легко доказываемого тождества

$$\frac{d}{d\alpha} \mathbf{M}^{\alpha} \varphi_n^* \Big|_{\alpha=\alpha_0} = \mathbf{M}^{\alpha_0} (\varphi_n^* - \alpha_0) \cdot \left[ \sum_{i=1}^n \frac{\partial}{\partial \alpha} \ln f(x_i; \alpha) \right]_{\alpha=\alpha_0}.$$

Из теоремы 1 и неравенства Крамера–Рао

$$\mathbf{M}^{\alpha_0} (\alpha_n^* - \alpha_0)^2 \geq \frac{\left(1 + \frac{d}{d\alpha} \mathbf{M}^{\alpha} \alpha_n^* \Big|_{\alpha=\alpha_0}\right)^2}{n k^2(\alpha_0)}$$

следует, что оценка наибольшего правдоподобия  $\varphi^*$  асимптотически наилучшая. Таким образом, имеет место следующая

ТЕОРЕМА 2. Если семейство законов распределения удовлетворяет условиям А, Б<sub>0</sub>–Б<sub>5</sub>, В<sub>1</sub>, В<sub>2</sub>, Г, то среди асимптотически несмещенных оценок, регулярных в смысле Крамера, существует класс эквивалентных друг другу асимптотически наилучших оценок.

Все они даются формулой  $\alpha^* = \varphi^* + \delta$ , где  $\varphi^*$  — оценка наибольшего правдоподобия,

$$\mathbf{M}^{\alpha} \delta_n = o\left(\frac{1}{\sqrt{n}}\right), \quad \mathbf{D}^{\alpha} \delta_n = o\left(\frac{1}{n}\right), \quad \frac{d}{d\alpha} \mathbf{M}^{\alpha} \delta_n = o(1).$$

Для каждой такой оценки  $\alpha^*$

$$\mathbf{M}^{\alpha_0} (\alpha_n^* - \alpha_0)^2 = \frac{1}{n k^2(\alpha_0)} + o\left(\frac{1}{n}\right).$$

В заключение автор считает приятным долгом выразить свою благодарность профессору Е.Б. Дынкину, под руководством которого была выполнена настоящая работа.

## СПИСОК ЛИТЕРАТУРЫ

1. *Каган А.М.* Некоторые свойства оценок максимального правдоподобия // Тр. VI Всес. сов. по теор. вер. и мат. статистике. — Вильнюс, 1962.
2. *Ченцов Н.Н.* Асимптотическая теория статистических оценок: Дипломная работа. — М.: Кафедра теории вероятностей МГУ, 1952.
3. *Ченцов Н.Н.* Об асимптотически наилучших статистических оценках параметра // Тр. III Всесоюзного математического съезда 1956 г. Т. 4. — М.: Изд-во АН СССР, 1959. — С. 71.
4. *Крамер Г.* Математические методы статистики. — М.: ИЛ, 1948.
5. *Колмогоров А.Н.* Несмещенные оценки // ИАН. 1950. Т. 14, № 4. С. 303–326.
6. *Wald A.* Note on consistency of the maximum likelihood estimate // Ann. Math. Statistics. 1949. V. 20. P. 595–601.
7. *Wolfowitz J.* On Wald's proof of the consistency of the maximum likelihood estimate // Ann. Math. Statistics. 1949. V. 20. P. 601–602.
8. *Diananda P.H.* Note on some properties of maximum likelihood estimates // Proc. Cambridge Philos. Soc. 1949. V. 45. P. 536–544.

## ОЦЕНКА НЕИЗВЕСТНОЙ ПЛОТНОСТИ РАСПРЕДЕЛЕНИЯ ПО НАБЛЮДЕНИЯМ<sup>\*)</sup>

*Представлено академиком М.В. Келдышем 21 V 1962*

Как известно (см. [1]), уклонение гистограммы от неизвестного графика плотности убывает, грубо говоря, как  $N^{-1/3}$ , где  $N$  — число независимых наблюдений, по которым построена гистограмма. В работе рассматривается класс методов оценки неизвестной плотности, обобщающих метод гистограмм и могущих дать большую точность. Эти способы оказались полезными автору при решении некоторых практических задач (см. [2]).

Пусть  $\xi$  — случайная величина (не обязательно числовая),  $\mu(dx)$  — мера в пространстве  $X$  значений  $\xi$ , функция точки;  $p(x) = dP/d\mu$  — искомая плотность распределения вероятностей  $P(\cdot)$  величины  $\xi$  относительно меры  $\mu$ .

Пусть на  $X$  задан вес  $r(x)$  и посредством скалярного произведения

$$(\varphi, \psi) = \int_X \varphi(x) \psi(x) r(x) \mu(dx) \quad (1)$$

задано гильбертово пространство  $L_2(r)$ , причем  $p(\cdot) \in L_2(r)$ .

Рассмотрим произвольное  $n$ -мерное подпространство  $E_n$  с ортонормированным базисом  $\{\varphi_{kn}(x)\}$ ,  $k = 1, \dots, n$ . Обозначим

$$\pi_n(x) = \sum_{k=1}^n a_{kn} \varphi_{kn}(x) = \sum_{k=1}^n (\varphi_{kn}, p) \varphi_{kn}(x)$$

<sup>\*)</sup> Докл. АН СССР. 1962. Т. 147, № 1. С. 45–48.

среднеквадратичную аппроксимацию  $p(x)$ , т.е. проекцию  $p(\cdot)$  на  $E_n$ . Коэффициенты  $a_{kn}$  могут быть истолкованы как

$$a_{kn} = (\varphi_{kn}, p) = \int \varphi_{kn}(x) r(x) dP = \mathbf{M}\alpha_k(\xi), \quad (2)$$

где  $\alpha_k(\xi) = \varphi_{kn}(\xi^i r(\xi))$ . Пусть  $\xi^{(1)}, \dots, \xi^{(N)}$  — независимые наблюдения величины  $\xi$ . Составим среднее  $\alpha_k^* = \frac{1}{N} = \sum_{k=1}^N \alpha_k(\xi^{(i)})$ , а также многочлен

$$\pi_{nN}^*(x) = \sum_{k=1}^n \alpha_k^* \varphi_{kn}(x) = \sum_{k=1}^n \left( \varphi_{kn}, \frac{d\Xi_N}{d\mu} \right) \varphi_{kn}(x), \quad (3)$$

т.е. проекцию на  $E_N$  “производной”  $d\Xi_N/d\mu$  по мере  $\mu(dx)$  эмпирической дискретной меры  $\Xi_N(dx)$  с атомами веса  $1/N$  в наблюдаемых точках  $\xi^{(i)}$ .

По теореме Пифагора

$$\|\pi^*(x) - p(x)\|^2 = \|\pi_n(x) - p(x)\|^2 + \sum_{k=1}^n [\alpha_k^* - a_{kn}]^2. \quad (4)$$

Таким образом, подобрав достаточно хорошо приближающее  $p(x)$  подпространство  $E_n$  и сделав затем достаточно большое число  $N$  наблюдений, мы можем получить по формуле (3) приближение к  $p(x)$  со сколь угодно  $L_2(r)$ -малой случайной ошибкой.

Хотя плотность  $p(x)$  не известна, у нас может иметься информация о том, что  $p(x)$  принадлежит множеству  $\Pi$  функций с известной скоростью аппроксимации (например, тригонометрическими полиномами). Пусть было сделано  $N$  независимых наблюдений  $\xi$ . Какую размерность  $n$  надо взять, чтобы использовать их наилучшим образом для оценки  $p(x)$ ? Скорость аппроксимации  $p(x)$ , заданной последовательностью подпространств  $E_n$ , предполагается известной. Ошибку  $\rho(\|\pi^* - p\|)$  в аппроксимации  $p(x)$  случайным многочленом  $\pi^*(x)$  измеряем или величиной  $\sqrt{\mathbf{M}\|\pi^* - p\|^2}$ , или доверительной оценкой  $Q_\delta(\|\pi^*(x) - p(x)\|)$  — квантилем порядка  $1 - \delta$  при соответствующем достаточно малом  $\delta$ .

ТЕОРЕМА 1а. Если известно, что  $p(\cdot) \in L_2(r)$  и

$$\|\pi_n(x) - p(x)\|^2 \leq Bg(Cn), \quad (5а)$$

а последовательность  $E_n$  такова, что

$$\int \left\{ \frac{1}{n} \sum_{k=1}^n [\varphi_{kn}(x)]^2 \right\} [r(x)]^2 p(x) \mu(dx) \leq H, \quad (6а)$$

где  $g(n) \downarrow 0$ ;  $B, C, H$  — положительные постоянные, то можно выбрать  $n$  так (грубо говоря,  $n_N \asymp \Gamma(N)$ ), чтобы

$$\rho(\|\pi^* - p\|) = O(\sqrt{N^{-1}\Gamma(N)}),$$

где  $\Gamma(y/g(y)) \equiv y$ .

ТЕОРЕМА 16. Если известно, что  $p(\cdot) \in L_2(r)$  и

$$\|\pi_n(x) - p(x)\|^2 \geq bg(cn), \quad (56)$$

$$\int \left\{ \frac{1}{n} \sum_{k=1}^n [\varphi_{kn}(x)]^2 \right\} [r(x)]^2 p(x) \mu(dx) \geq h, \quad (66)$$

где  $b, c, h$  — некоторые положительные постоянные, а при измерении уклонения по квантилю также, что при достаточно малых  $\delta$

$$Q_\delta \left( \sum_{k=1}^n [\alpha_k^* - a_{kn}]^2 \right) \geq lM \left( \sum_{k=1}^n [\alpha_k^* - a_{kn}]^2 \right),$$

то при любом выборе  $n$

$$\rho(\|\pi^* - p\|) = \Omega(\sqrt{N^{-1}\Gamma(N)}).$$

Таким образом, если условия теорем выполнены, норма ошибки  $\|\pi^* - p\|$  в лучшем случае имеет порядок  $\sqrt{N^{-1}\Gamma(N)}$ . Для доказательства теорем оценивается сверху и снизу минимум по  $n$  математического ожидания или квантиля суммы (4). Первое слагаемое убывает как  $g(n)$ , второе слагаемое ведет себя как  $n/N$ , а их сумма будет почти минимальной при равенстве слагаемых. Обе эти теоремы могут быть перенесены на случай переменного веса  $r_n(x)$ .

СЛЕДСТВИЕ 1. Пусть  $\pi_{nN}^*(x)$  — гистограмма случайной величины  $\xi$ ,  $a \leq \xi \leq b$ , с плотностью распределения  $p(x)$ , построенная по  $N$  независимым наблюдениям.  $L_2$ -уклонение  $\pi^*(x)$  от  $p(x)$  в лучшем случае (в частности, при  $n \asymp \sqrt[3]{N}$ , равных по длине интервалах группировки) имеет по вероятности порядок  $N^{-1/3}$ . (При этом  $p'(x)$  предполагается непрерывной и не равной нулю тождественно.)

СЛЕДСТВИЕ 2.  $L_2$ -уклонение гистограммы  $\pi_{nN}^*(\vec{x})$  векторной случайной величины  $\vec{\xi} = \{\xi_m\}$ ,  $a_m \leq \xi_m \leq b_m$ ;  $m = 1, \dots, s$ , в лучшем случае имеет по вероятности порядок  $N^{-1/(s+2)}$ .

СЛЕДСТВИЕ 3. Если  $\{\varphi_k(x)\}$  — ортонормированный базис в  $L_2(r)$  и  $|(\varphi_k, p)| \leq Ak^{-m}$ ,  $\int [\varphi_k(x)]^2 p(x) \mu(dx) \leq H$ , то при выборе  $n \asymp \sqrt[2m]{N}$  уклонение  $\|\pi^* - p\|$  имеет по вероятности порядок  $N^{-1/2+1/(4m)}$ .

Поскольку скорость аппроксимации  $p(x)$  отрезком ряда  $\sum a_k \varphi_k(x)$  заранее может быть не известна, то можно подбирать  $n$  по результатам наблюдений, ограничиваясь в (3) только такими слагаемыми, у которых коэффициент  $\alpha_{kN}^*$  существенно больше своей экспериментальной среднеквадратичной ошибки.

Выше рассматривалась аппроксимация  $p(x)$  в метрике  $L_2$ . В других метриках теоремы 1а и 1б могут дать оценки сверху или снизу. В частности, получается, что максимум модуля уклонения гистограммы от плотности есть величина  $\Omega(N^{-1/3})$ . Из результатов Н.В. Смирнова [1] (см. также [3]) вытекает, что эту оценку можно уточнить только на логарифмический множитель. В отличие от гистограммы, при  $\varphi_{kn}(x)$  общего вида многочлен  $\pi_{kN}^*(x)$  может оказаться неположительным. Чтобы избежать этого, можно пытаться его подправить, пожив, например,  $\tilde{\pi}_{nN}^*(x) = \sum_{k=1}^n \gamma_{kn} \alpha_{kn}^* \varphi_{kn}(x)$ , где  $\gamma_{kn}$  — некоторые множители.

Заранее очевидно, что оценка  $p(x)$  по формуле (3) целесообразна далеко не всегда. Однако если нужна оценка, пригодная для любой неизвестной плотности  $p(x)$  из существенно бесконечномерного множества  $\Pi$  плотностей, то оценка (3) может оказаться близкой к оптимальной (точнее, с ошибкой, лишь в конечное число раз большей, чем у оптимальной оценки). Рассмотрим два ограничения на множество  $\Pi$  и вес  $r(x)$ :

$$\sup_{p \in \Pi, x \in X} p(x) r(x) \leq A, \quad (7a)$$

$$\inf_{p \in \Pi, x \in X} p(x) r(x) \geq a. \quad (7б)$$

Если условия (7a) и (7б) выполнены одновременно, то все плотности  $p(\cdot) \in \Pi$  равномерно “лишшицируемы” друг относительно друга:

$$\frac{a}{A} \leq \inf_{p, q, x} \frac{p(x)}{q(x)} \leq \sup_{p, q, x} \frac{p(x)}{q(x)} \leq \frac{A}{a}. \quad (8)$$

Обозначим через  $d_n(\Pi)$   $n$ -мерный поперечник  $\Pi$  (см. [4]), а через  $r_n(\Pi)$  — радиус максимальной вписанной в  $\Pi$   $n$ -мерной сферы.

ТЕОРЕМА 2a. Если условие (7a) выполнено и

$$[d_n(\Pi)]^2 \leq Bg(Cn), \quad (9a)$$

то можно выбрать такую последовательность подпространства  $E_n$  и такую зависимость  $n$  от  $N$ , что

$$\sup_{p \in \Pi} \rho(\|\pi_{nN}^* - p\|) = O(\sqrt{N^{-1}\Gamma(N)}).$$

ТЕОРЕМА 2б. Если условие (7б) выполнено и

$$[r_n(\Pi)]^2 \geq bg(cn), \quad (9б)$$

то для любой оценки  $\pi_N^*(x; \xi^{(1)}, \dots, \xi^{(N)})$ , каков бы ни был способ ее построения,

$$\sup_{p \in \Pi} \rho(\|\pi_N^* - p\|) = \Omega(\sqrt{N^{-1}\Gamma(N)}).$$

Теорема 2a легко следует из теоремы 1a. Идея доказательства теоремы 2б состоит в том, что, с одной стороны,  $M\|\pi^* - p\|^2 \geq \|M\pi^* - p\|^2$ , а с другой — по неравенству Крамера–Рао (см. [5]) для любого  $n$ -мерного подпространства  $E_n \ni p$   $M\|\pi^* - p\|^2 \geq a \frac{n}{N} \left[ \frac{1}{n} \operatorname{div} M\pi_n^* \right]^2$ , где  $\pi_n^*$  — проекция  $\pi^*$  на  $E_n$ .

Если  $\Pi$  обладает свойством (8), а каждый вес  $r(x) \in R$  обладает свойствами (7a) и (7б), то все метрики  $L_2(r)$  эквивалентны на  $\Pi$ . При этих условиях теорему 2a можно усилить, поставив в ней  $\sup_{p \in \Pi} \sup_{r \in R} \rho(\|\pi_{nN}^* - p\|)$ , а в теореме 2б  $\sup_{p \in \Pi} \inf_{r \in R} \rho(\|\pi_N^* - p\|)$ .

Наиболее важный пример семейства весов  $r(x) = 1/p(x)$ , в каждой точке  $p \in \Pi$  свое скалярное произведение, своя метрика, так что  $\Pi$  можно трактовать

как риманово многообразие бесконечной размерности. В этом случае качество аппроксимации определяется средним квадратом относительной ошибки. Следует, однако, отметить, что для распределений, заданных на всей действительной прямой или на всем эвклидовом пространстве, условия (7а) и (7б), тем более (8), оказываются чересчур ограничительными.

В заключение автор считает приятным долгом выразить благодарность Н.В. Смирнову за постоянное внимание и ценную дискуссию.

#### СПИСОК ЛИТЕРАТУРЫ

1. Смирнов Н.В. // Докл. АН СССР. 1950. Т. 74, № 2. С. 189.
2. Гельфанд И.М., Фейнберг С.М., Фролов А.С., Ченцов Н.Н. // Тр. II Международн. конф. ООН по мирн. использов. атомн. энергии. Т. 2. — М., 1959. — С. 628.
3. Гливенко В.И. Курс теории вероятностей. — М., 1939.
4. Тихомиров В.М. // Успехи матем. наук. 1960. Т. 15, вып. 3. С. 81.
5. Крамер Г. Математические методы статистики. — М.: ИЛ, 1948.

## ОБ ОЦЕНКЕ НЕИЗВЕСТНОГО СРЕДНЕГО МНОГОМЕРНОГО НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ<sup>\*)</sup>

### § 1. Введение. Постановка задачи оптимального оценивания по Гауссу и Вальду

Пусть  $\mathfrak{X}$  — семейство  $s$ -мерных нормальных законов  $\Phi_{\mathbf{a}}$  с единичной матрицей ковариаций и векторным параметром — средним значением  $\mathbf{a} = (a_1, \dots, a_s)$ . Соответствующие плотности распределений вероятностей образуют семейство

$$p(\mathbf{x}; \mathbf{a}) = (2\pi)^{-s/2} \exp \left\{ -\frac{1}{2} \|\mathbf{x} - \mathbf{a}\|^2 \right\}, \quad (1)$$

где  $\|\mathbf{x}\|^2 = x_1^2 + \dots + x_s^2$ .

Дано, что наблюдения  $\xi^1, \dots, \xi^{(N)}$  независимы и одинаково распределены по нормальному закону с плотностью (1), однако значение параметра  $\mathbf{a}$  этого закона не известно наблюдателю, и он должен оценить его с помощью своих наблюдений. Дано также, что оценка искомого  $\mathbf{a}$  величиной  $\alpha$  приведет к потерям  $L(\alpha, \mathbf{a}) = \|\alpha - \mathbf{a}\|^2$ . Априорное качество оценки характеризуется функцией риска  $R(\mathbf{a}) = \mathbf{M}_{\mathbf{a}} \|\alpha - \mathbf{a}\|^2$  — математическим ожиданием потерь в предположении, что наблюдения распределены по закону  $\Phi_{\mathbf{a}}$ . Проблема состоит в том, чтобы построить по наблюдениям оптимальную оценку, минимизирующую риск<sup>1)</sup>.

<sup>\*)</sup> Теория вероятн. и ее примен. 1967. Т. 12, вып. 4. С. 619–633.

<sup>1)</sup> Или, более строго, наименее рискованную оценку (при соответствующем точном определении величины рискованности оценки через ее функцию риска  $R(\cdot)$ ; см. § 2).

Исходное семейство плотностей (1) инвариантно относительно группы евклидовых движений пространства параметров. Предложенная Гауссом (см. [1, 2]) функция потерь  $\|\alpha - \mathbf{a}\|^2$  также инвариантна относительно этой группы. В силу указанной однородности семейство  $\mathfrak{X}$  является излюбленным простейшим объектом для демонстрации принципов и методов статистического оценивания. В настоящей работе мы хотим на его примере обсудить возможные точные постановки задачи построения оптимальных оценок и выяснить их асимптотические свойства. В своих рассуждениях мы постараемся использовать только такие инвариантные свойства и характеристики семейства  $\mathfrak{X}$ , которые у других семейств являются свойствами и характеристиками, инвариантными в категории статистических задач (см. [3–6]). Такой подход позволит нам предложить “естественные” постановки задачи асимптотически оптимального оценивания для гладких семейств распределений самого общего вида и предугадать точную формулировку ее решения (см. [6]). В то же время мы сможем достаточно кратко изложить аккуратные доказательства своих предложений. В общем же случае они были бы перегружены техническими деталями и требовали бы введения новых понятий и определений, вроде инвариантности в категории и гладкости поверхности в функциональном пространстве.

## § 2. Характеристика качества решающего правила. Оптимальная инвариантная оценка

Естественно в первую очередь искать оптимум среди инвариантных оценок. Простейшей оценкой является среднее арифметическое наблюдений

$$\eta^{(N)} = \frac{1}{N} [\xi^{(1)} + \dots + \xi^{(N)}]. \quad (2)$$

Эта оценка эффективна: она несмещена,  $\mathbf{M}_\mathbf{a}\eta = \mathbf{a}$ ; ее компоненты некоррелированы,  $\mathbf{M}_\mathbf{a}\eta_j\eta_k = 0$  при  $j \neq k$ , а их дисперсии  $\mathbf{D}_\mathbf{a}\eta_j^{(N)}$  как функции параметра  $\mathbf{a}$  закона распределения  $\Phi_\mathbf{a}$  постоянны и всюду достигают минимума, разрешаемого (для несмещенных оценок) неравенством информации  $\mathbf{D}_\mathbf{a}\alpha_j^{(N)} \geq 1/N$  (см. [7, 8]). Отсюда следует, что оценка  $\eta^{(N)}$  минимизирует функцию риска  $R(\mathbf{a})$  при всех значениях параметра  $\mathbf{a}$ :

$$\mathbf{M}_\mathbf{a}\|\eta^{(N)} - \mathbf{a}\|^2 = \frac{s}{N}, \quad (3)$$

и является поэтому оптимальной в классе всех несмещенных оценок. Притом она является единственной эффективной (см. [7]), а следовательно, и единственной оптимальной оценкой в том смысле, что всякая несмещенная оценка  $\alpha^{(N)}$  с функцией риска  $R(\mathbf{a}) = s/N$  совпадает с  $\eta^{(N)}$  с вероятностью единица.

У любой инвариантной оценки  $\mathbf{a}$  вектор смещения постоянен:  $\mathbf{M}_\mathbf{a}\alpha - \mathbf{a} = \mathbf{c}$ , и

$$\mathbf{M}_\mathbf{a}\|\alpha - \mathbf{a}\|^2 = \sum_j \mathbf{D}_\mathbf{a}\alpha_j + \|\mathbf{M}_\mathbf{a}\alpha - \mathbf{a}\|^2.$$

Поэтому, несмещенная оценка  $\beta = \alpha - \mathbf{c}$ ,  $\mathbf{D}_\mathbf{a}\beta_j = \mathbf{D}_\mathbf{a}\alpha_j$ , лучше оценки  $\alpha$ , а  $\eta^{(N)}$  лучше всякой другой несмещенной оценки. Следовательно, оценка  $\eta^{(N)}$  является оптимальной и в классе всех инвариантных оценок.

Среди всех без ограничения оценок дело обстоит, как известно, значительно сложнее. Во-первых, заведомо не существует абсолютно наилучших оценок. В самом деле, оценка вида  $\alpha = \mathbf{a}^0$  абсолютно точна при  $\mathbf{a} = \mathbf{a}^0$  (хотя и неприемлема при больших  $\|\mathbf{a} - \mathbf{a}^0\|$ ), а у семейства  $\mathfrak{R}$  в силу взаимной абсолютной непрерывности законов  $\Phi_{\mathbf{a}}$  нет оценок, абсолютно точных хотя бы для двух (не говоря уже обо всех) значениях параметра  $\mathbf{a}$ . Поэтому уже сама постановка задачи нуждается в уточнении: надо указать, что понимается под минимизацией функции риска. Во-вторых, как показал Стейн [9], при  $s \geq 3$  существуют инвариантные смещенные оценки, абсолютно лучшие, чем среднее  $\eta^{(N)}$ , т.е. лучшие при всех значениях параметра  $\mathbf{a}$ . Например (см. [10, 11]), такой оценкой является

$$\zeta^{(N)} = \left(1 - \frac{s-2}{\|\eta^{(N)}\|^2}\right) \eta^{(N)}. \quad (4)$$

Поэтому оптимальные оценки заведомо должны быть достаточно сложными и существенно зависеть от точной формулировки принципа оптимальности.

Чтобы иметь возможность сравнивать между собой любые два решающих правила  $\Pi_1$  и  $\Pi_2$ , надо условиться о способе сравнения соответствующих функций риска  $R_1(\mathbf{a})$  и  $R_2(\mathbf{a})$ , измеряющих их качество. Для этого требуется дополнительно задать на множестве функций риска некоторый функционал  $Q(\Pi) = Q[R_{\Pi}^*(\cdot)]$ . Естественно, что этот функционал должен быть монотонным:

$$Q[R_1(\cdot)] \geq Q[R_2(\cdot)], \quad \text{если } R_1(\mathbf{a}) \geq R_2(\mathbf{a}) \quad \text{при всех } \mathbf{a}. \quad (5)$$

Далее, поскольку монотонное преобразование шкалы  $Q$  не изменяет результатов сравнений, не особенно теряя в общности, потребуем также, чтобы

$$Q[\lambda \cdot 1(\mathbf{a})] = \lambda, \quad \text{где } \lambda \geq 0 \quad \text{и} \quad 1(\mathbf{a}) = 1 \quad \text{при всех } \mathbf{a}. \quad (6)$$

Такой показатель  $Q(\Pi)$  можно в общем случае произвольной функции потерь  $L(\alpha, \mathbf{a})$  назвать рискованностью правила, но при естественной гауссовской функции потерь  $\|\alpha - \mathbf{a}\|^2$  целесообразнее, пожалуй, слово “неопределенность”, хотя оно уже употребляется в квантовой механике в близком, но не идентичном смысле.

### 3. Вспомогательные леммы

Проще всего измерять качество оценки максимумом (или, точнее, верхней гранью) риска. Наша ближайшая цель — показать, что среднее арифметическое  $\eta^{(N)}$  является минимаксной, т.е. оптимальной в этом смысле оценкой. Сформулируем сперва ряд лемм. Их доказательства будут опущены ввиду элементарности.

ЛЕММА 1. Пусть  $\Pi\{d\alpha|\xi^{(1)}, \dots, \xi^{(N)}\}$  — решающее правило, задающее распределение рандомизированной оценки  $\alpha$  неизвестного параметра  $\mathbf{a}$  по наблюдениям  $\xi^{(1)}, \dots, \xi^{(N)}$ , и пусть

$$\bar{\alpha} = \bar{\alpha}(\xi^{(1)}, \dots, \xi^{(N)}) = \int \alpha \Pi\{d\alpha|\xi^{(1)}, \dots, \xi^{(N)}\}.$$

Тогда оценка  $\bar{\alpha}$  абсолютно лучше оценки  $\alpha$ , т.е. при всех  $\mathbf{a}$

$$R_{\Pi}(\mathbf{a}) = \mathbf{M}_{\mathbf{a}}\|\alpha - \mathbf{a}\|^2 \geq \mathbf{M}_{\mathbf{a}}\|\bar{\alpha} - \mathbf{a}\|^2,$$

где слева математическое ожидание берется по всем исходам эксперимента.

Таким образом, отыскивая оптимальные оценки, можно ограничиться только детерминированными функциями наблюдений. Заметим, что этот вывод справедлив для любой выпуклой по  $\alpha$  функции потерь  $L(\alpha, \mathbf{a})$ .

**ЛЕММА 2.** Пусть  $\alpha = \{\alpha_j(\xi^{(1)}, \dots, \xi^{(N)})\}_{j=1}^s$  — вектор-функция независимых наблюдений  $\xi^{(i)}$ , распределенных по закону  $\Phi_{\mathbf{a}}$  каждое, и при всех  $\mathbf{a}$  из некоторого множества  $G$  существует  $\mathbf{M}_{\mathbf{a}}\|\alpha\|^2 < \infty$ .

Тогда в каждой точке  $\mathbf{a}$  выпуклой оболочки  $(G)$  множества  $G$  математическое ожидание  $\mathbf{M}_{\mathbf{a}}\|\alpha\|^2$  конечно и как функция  $\mathbf{a}$  непрерывно внутри открытого ядра  $O$  этой оболочки  $(G)$ . В каждой точке  $(G)$  существует интеграл

$$\mathbf{M}_{\mathbf{a}}\alpha = \int \dots \int \alpha(x^{(1)}, \dots, x^{(N)}) \prod_{i=1}^N [p(x^{(i)}; \mathbf{a}) dx^{(i)}]. \quad (7)$$

В области  $O$  он является аналитической функцией от  $\mathbf{a}$ , а его производные по параметрам  $a_1, \dots, a_s$  можно вычислять дифференцированием подынтегральной функции.

Аналогичное утверждение справедливо для любых экспоненциальных семейств распределений (ср. [12, гл. II, § 7]).

**ЛЕММА 3.** Пусть  $p(\omega; \mathbf{t})$  — гладкое семейство плотностей распределений вероятностей,

$$\int p(\omega; \mathbf{t}) d\mu = 1, \quad (8)$$

$\mathbf{t} = (t_1, \dots, t_m)$  — параметр,  $\tau = (\tau_1(\omega), \dots, \tau_m(\omega))$  — векторная статистика, и

$$\mathbf{M}_{\mathbf{t}}\tau(\omega) = \int \tau(\omega) p(\omega; \mathbf{t}) d\mu = \mathbf{T}(\mathbf{t}), \quad (9)$$

причем интегралы (8) и (9) в некоторой точке  $\mathbf{t} = \theta$  можно дифференцировать по параметру  $\mathbf{t}$ , производя дифференцирование под знаком интеграла.

Тогда при всех  $j = 1, \dots, m$

$$\mathbf{M}_{\theta}[\tau_j(\omega) - \theta_j]^2 \geq [T_j(\theta) - \theta_j]^2 + \frac{(\partial T_j / \partial t_j)_{\mathbf{t}=\theta}^2}{\mathbf{M}_{\theta}[\partial \ln p(\omega; \mathbf{t}) / \partial t_j]_{\mathbf{t}=\theta}^2}. \quad (10)$$

Утверждение леммы есть частный случай неравенства информации (см., например, [7, 8]).

**ЛЕММА 4.** Пусть  $\alpha = \alpha(\xi^{(1)}, \dots, \xi^{(N)})$  — оценка параметра  $\mathbf{a}$  закона  $\Phi_{\mathbf{a}}$  по наблюдениям  $\xi^{(1)}, \dots, \xi^{(N)}$  и  $\mathbf{A}(\mathbf{a}) = \mathbf{M}_{\mathbf{a}}\alpha$ .

Тогда значения параметра  $\mathbf{a}$ , при которых функция риска  $R_{\alpha(\cdot)}(\mathbf{a}) = \mathbf{M}_{\mathbf{a}}\|\alpha - \mathbf{a}\|^2$  конечна, образуют выпуклое множество, причем во всех его внутренних точках выполнено неравенство

$$\mathbf{M}_{\mathbf{a}}\|\alpha - \mathbf{a}\|^2 \geq \|\mathbf{A}(\mathbf{a}) - \mathbf{a}\|^2 + \frac{1}{N} \sum_{j=1}^s \left( \frac{\partial A_j}{\partial a_j} \right)^2. \quad (11)$$

Доказательство леммы немедленно вытекает из лемм 2 и 3. Неравенство (11) можно усиливать; например, если вместо (10) воспользоваться уточненным неравенством информации (см. [13]), можно к правой части прибавить еще член

$$\frac{1}{2N^2} \sum \left( \frac{\partial^2 A_j}{\partial a_j^2} \right)^2.$$

#### § 4. Оценка усредненного риска. Интегральные следствия неравенства информации

ЛЕММА 5. Пусть  $\mathbf{w}(\mathbf{z})$  — непрерывно дифференцируемое отображение открытого  $s$ -мерного куба  $C_r = \{\mathbf{z} : |z_j| < r, j = 1, \dots, s\}$  во все  $s$ -мерное пространство.

Тогда для среднего  $\Psi_r^{(\varkappa)}(\mathbf{w})$  по кубу

$$\Psi_r^{(\varkappa)}(\mathbf{w}) = \frac{1}{(2r)^s} \int_{-r}^r \cdots \int_{-r}^r \left[ \frac{1}{\varkappa^2} \sum_{j=1}^s \left( \frac{\partial w_j}{\partial z_j} \right)^2 + \sum_{j=1}^s (w_j - z_j)^2 \right] dz_1 \dots dz_s$$

справедливо неравенство

$$\Psi_r^{(\varkappa)}(\mathbf{w}) \geq \frac{s}{\varkappa^2} \left[ 1 - \frac{1}{\varkappa r} \operatorname{th} \varkappa r \right]. \quad (12)$$

Наметим доказательство леммы. Система уравнений Эйлера для экстремали  $\mathbf{v}$  функционала  $\Psi_r^{(\varkappa)}(\mathbf{w})$  распадается:

$$-\frac{1}{\varkappa^2} \frac{\partial^2 v_j}{\partial z_j^2} + (v_j - z_j) = 0, \quad \frac{\partial v_j}{\partial z_j} \Big|_{z_j = \pm r} = 0, \quad j = 1, \dots, s. \quad (13)$$

Обозначив  $v_j - z_j = u_j$ , приведем каждое из этих уравнений к виду

$$\frac{1}{\varkappa^2} \frac{d^2 u}{dz^2} - u = 0, \quad \frac{du}{dz} \Big|_{z = \pm r} = -1.$$

Решением последнего уравнения будет  $u = \frac{1}{\varkappa} \frac{\operatorname{sh} \varkappa z}{\operatorname{ch} \varkappa r}$ . Предположим теперь, что отображение  $\mathbf{w}(\mathbf{z})$  непрерывно, включая границу  $C_r$ . Функция  $u(z)$  аналитическая, следовательно, экстремаль  $\mathbf{v}(\mathbf{z})$ ,  $v_j(z) = z_j + u(z_j)$ ,  $j = 1, \dots, s$ , гладкая и реализует абсолютный минимум квадратичного функционала  $\Psi_r^{(\varkappa)}(\mathbf{w})$ , приводя его к сумме положительных квадратичных форм  $\Psi_r^{(\varkappa)}(\mathbf{w}) = \Psi_r^{(\varkappa)}(\mathbf{v}) + \psi(\mathbf{w} - \mathbf{v})$ , где  $\psi(\mathbf{w} - \mathbf{v})$  — среднее по кубу от

$$\frac{1}{\varkappa^2} \sum_j \left[ \frac{\partial (w_j - v_j)}{\partial z_j} \right]^2 + \sum_j (w_j - v_j)^2.$$

Как обычно, после интегрирования по частям, законного по предположенной непрерывности  $\mathbf{w}$ , линейные члены уничтожаются в силу уравнений Эйлера (13).

Интегралы от функции  $u$  и ее производной, встречающиеся при вычислении  $\Psi_r^{(\varkappa)}(\mathbf{v})$ , легко берутся, так что нетрудно проверить, что неравенство (12) есть неравенство  $\Psi_r^{(\varkappa)}(\mathbf{w}) \geq \Psi_r^{(\varkappa)}(\mathbf{v})$ . Перейдем к общему случаю. Очевидно, для интегралов по  $C_r$  и любому вложенному кубу  $C_{r-\varepsilon}$  справедливы неравенства

$$(2r)^s \Psi_r^{(\varkappa)}(\mathbf{w}) \geq (2r - 2\varepsilon)^s \Psi_r^{(\varkappa)}(\mathbf{w}) \geq s \frac{(2r - 2\varepsilon)^s}{\varkappa^2} \left[ 1 - \frac{1}{\varkappa(r - \varepsilon)} \operatorname{th} \varkappa(r - \varepsilon) \right],$$

так как в силу условия леммы отображение  $\mathbf{w}(\mathbf{z})$  непрерывно в  $C_{r-\varepsilon}$ . Поделив левую и правую части неравенства на  $(2r)^s$  и устремив  $\varepsilon$  к нулю, получаем искомое неравенство (12).

**ТЕОРЕМА 1.** Пусть  $\Pi$  — решающее правило для оценки  $\alpha$  неизвестного параметра  $\mathbf{a}$  закона  $\Phi_{\mathbf{a}} \in \mathfrak{R}$  по  $N$  независимым наблюдениям. Пусть, далее,  $R_{\Pi}(\mathbf{a})$  — риск, отвечающий функции потерь  $\|\alpha - \mathbf{a}\|^2$ .

Тогда для осреднения  $\mathfrak{R}_{\Pi}(C_r)$  риска по любому кубу  $C_r(b) = \{\mathbf{a} : |a_j - b_j| < r, j = 1, \dots, s\}$  в пространстве параметров справедливо неравенство

$$\mathfrak{R}_{\Pi}(C_r) = \frac{1}{(2r)^s} \int_{b_1-r}^{b_1+r} \dots \int_{b_s-r}^{b_s+r} R_{\Pi}(\mathbf{a}) da_1 \dots da_s \geq \frac{s}{N} \left[ 1 - \frac{1}{r\sqrt{N}} \operatorname{th} r\sqrt{N} \right]. \quad (14)$$

**Доказательство.** В силу леммы 1 достаточно проверить неравенство (14) только для детерминированных решающих правил. Пусть  $\alpha = \alpha(\xi^{(1)}, \dots, \xi^{(N)})$  — соответствующая оценка,  $\mathbf{A}(\mathbf{a}) = \mathbf{M}_{\mathbf{a}}\alpha$  — ее математическое ожидание,  $R(\mathbf{a}) = \mathbf{M}_{\mathbf{a}}\|\alpha - \mathbf{a}\|^2$  — соответствующая функция риска, а вектор  $\mathbf{a}$  — параметр закона  $\Phi_{\mathbf{a}}$  распределения наблюдений  $\xi^{(i)}$ . Из леммы 2 вытекает, что множество  $G$  значений  $\mathbf{a}$ , где  $R(\mathbf{a})$  конечно, является выпуклым множеством. Его пересечение  $G \cap C_r$  с кубом также выпукло. Следовательно, и множество  $G \cap C_r$ , и дополнительное множество  $H = C_r - (G \cap C_r)$  точек, где  $R(\mathbf{a})$  бесконечно, измеримы по Жордану. Внутри  $G$  функция риска положительна и непрерывна, поэтому ее интеграл по  $C_r$  существует, хотя и может оказаться бесконечным. Если множество  $H$  имеет ненулевой объем, то интеграл заведомо бесконечен, и неравенство (14) тривиально. Так что должен быть разобран лишь случай, когда  $H$  имеет нулевой объем и объем открытого куба  $C_r$  совпадает с объемом его выпуклого подмножества  $G \cap C_r$ . Но последнее может быть лишь когда  $H$  пусто (ибо от любой шаровой окрестности внутренней точки куба, не принадлежащей данному выпуклому множеству, т.е.  $G \cap C_r$ , можно отсечь полшара, т.е. множество положительного объема, также лежащего вне  $G \cap C_r$ ). Таким образом, мы вправе считать, что всюду внутри куба  $C_r$  функция риска конечна, и, следовательно, согласно лемме 4 она непрерывна и удовлетворяет в каждой точке неравенству (11), где правая часть неравенства также непрерывна. Проинтегрируем обе части (11) по всему кубу  $C_r(\mathbf{b})$  и разделим на его объем. Поскольку после переноса начала координат в точку  $\mathbf{b}$  условия леммы 5 выполнены, то

$$\frac{1}{(2r)^s} \int_{C_r(\mathbf{b})} \dots \int R_{\Pi}(\mathbf{a}) da_1 \dots da_s \geq \Psi_r^{(\sqrt{N})}(\mathbf{A}(\cdot)) \geq \frac{s}{N} \left( 1 - \frac{1}{r\sqrt{N}} \operatorname{th} r\sqrt{N} \right),$$

что и требовалось. Так как неравенство (11) было неточным, то (14) также можно усилить.

**СЛЕДСТВИЕ.** Среднее арифметическое  $\eta^{(N)}$  является минимаксной оценкой неизвестного параметра  $\mathbf{a}$ , т.е.  $\eta^{(N)}$  оптимально в этом смысле.

В самом деле, верхняя грань функции риска не меньше ее среднего значения по любому кубу, а последнее при больших  $r$  сколь угодно близко к  $s/N$  — величине риска  $\eta^{(N)}$ .

Ранее этот результат был установлен лишь при  $s = 1, 2$ . Именно, в работах [14, 15, 16] для  $s = 1$  и в работе [9] для  $s = 2$  было доказано, что среднее  $\eta^{(N)}$  является минимаксной и притом единственной минимаксной оценкой. При  $s \geq 3$ , как следует из примера Стейна (4), существуют суперэффективные минимаксные оценки с  $R(\mathbf{a}) < s/N$  при каждом  $\mathbf{a}$ . Нетрудно видеть, однако, что они мало где могут быть существенно лучше среднего  $\eta^{(N)}$ . Именно (см. [6]), для любой суперэффективной в указанном смысле оценки  $\alpha$  параметра  $\mathbf{a}$ , построенной по  $N$  наблюдениям, и для любого  $\varepsilon > 0$  доля множества

$$A_\varepsilon = \{\mathbf{a} : M_{\mathbf{a}} \|\alpha - \mathbf{a}\|^2 \leq s/N - \varepsilon\}$$

во всем пространстве равна нулю. Более точно, для любого куба  $C_r(\mathbf{b})$

$$\frac{V\{A_\varepsilon \cap C_r\}}{V\{C_r\}} \leq \frac{1}{\varepsilon} \cdot \frac{1}{r\sqrt{N}} \operatorname{th} r\sqrt{N},$$

где  $V$  — лебегов объем в  $s$ -мерном евклидовом пространстве параметров.

## § 5. Асимптотическая трудность задачи статистического оценивания

**ТЕОРЕМА 2.** Пусть качество оценки неизвестного параметра  $\mathbf{a}$  закона  $\Phi_{\mathbf{a}} \in \mathfrak{X}$  распределения наблюдений измеряется произвольным монотонным (см. (5)) и калиброванным (см. (6)) функционалом  $Q[R(\cdot)]$  от гауссовской функции риска, обладающим, кроме того, свойствами, во-первых, выпуклости и, во-вторых, инвариантности относительно евклидовых движений  $g$  пространства параметров

$$Q[R_1(\cdot)] = Q[R_2(\cdot)] \quad \text{при} \quad R_2(\mathbf{a}) = R_1(\mathbf{g}\mathbf{a}).$$

Тогда каково бы ни было решающее правило  $\Pi$ , построенное по  $N$  наблюдениям, его неопределенность

$$Q(\Pi_N) \geq \frac{s}{N}. \quad (15)$$

**Доказательство.** При любом  $r$  в силу последовательно калиброванности  $Q$ , монотонности  $Q$  и неравенства (14), выпуклости  $Q$ , инвариантности  $Q$  справедлива цепочка соотношений

$$\begin{aligned} \frac{s}{N} \left(1 - \frac{1}{r\sqrt{N}} \operatorname{th} r\sqrt{N}\right) &= Q \left[ \frac{s}{N} \left(1 - \frac{1}{r\sqrt{N}} \operatorname{th} r\sqrt{N}\right) \right] \leq \\ &\leq Q \left[ \frac{1}{(2r)^s} \int_{-r}^r \dots \int_{-r}^r R(\mathbf{a} - \mathbf{g}) \, dg_1 \dots dg_s \right] \leq \\ &\leq \frac{1}{(2r)^s} \int_{-r}^r \dots \int_{-r}^r Q[R(\mathbf{a} - \mathbf{g})] \, dg_1 \dots dg_s = Q[R(\mathbf{a})]. \end{aligned}$$

Отбросив промежуточные соотношения и устремив  $r$  к бесконечности, получим искомое.

“Красивая” формулировка (15) выглядит скорее забавно. И нам не стоило бы столь тщательно доказывать такой почти<sup>2)</sup> очевидный заранее факт, если бы аналогичный закон не был справедлив, правда, в ослабленной, асимптотической форме, при весьма широких предположениях (см. [6]). Уговоримся назвать асимптотической трудностью задачи статистического оценивания предел

$$\lim_{N \rightarrow \infty} N \cdot \inf_{\Pi} Q(\Pi_N). \quad (16)$$

В условиях теоремы 2 справедливо утверждение:

*трудность задачи статистического оценивания равна ее размерности.*

Это утверждение, являющееся интегральным следствием неравенства информации, имеет, как оказывается, характер закона при достаточно широких предположениях о гладкости априорного семейства распределений, конкретном выборе инвариантного способа измерения случайных ошибок и дополнительной априорной информации. Наша дальнейшая цель — обсудить границы его применимости.

### § 6. Повышение точности оценок при учете дополнительной априорной информации

Когда наблюдатель располагает дополнительной априорной информацией о распределении  $\Phi_{\mathbf{a}}$  вероятности наблюдений, его априорная оценка качества решающего правила естественным образом видоизменяется.

Пусть наблюдателю заранее известно, что параметр  $\mathbf{a}$  распределения  $\Phi_{\mathbf{a}}$  принадлежит к данному множеству  $F$ . Тогда априори возможный максимальный риск данного способа  $\Pi$  оценивания характеризуется величиной

$$Q_F(\Pi) = \sup_{\mathbf{a} \in F} R_{\Pi}(\mathbf{a}), \quad (17)$$

что может быть, вообще говоря, существенно меньше, чем при отсутствии дополнительной информации, и никогда не больше. Указанная дополнительная информация может быть также использована наблюдателем для улучшения своих решающих правил. Когда множество  $F$  выпукло, можно уменьшить рискованность  $Q_F$  способа оценивания общего вида, заменяя всякий раз оценку  $\alpha$  ее “проекцией”  $\beta_F(\alpha)$  на  $F$ . Это вытекает из следующего предложения.

**ЛЕММА 6.** Пусть  $G$  — выпуклое множество в евклидовом пространстве.

Тогда для любой точки  $\alpha$  этого пространства существует и единственна  $\beta_G(\alpha)$  ближайшая к  $\alpha$  точка замыкания  $G$ . Точка  $\beta_G(\alpha)$  есть точка границы  $G$ , когда  $\alpha \notin G$ , и  $\beta_G(\alpha) = \alpha$ , когда  $\alpha \in G$ . При этом для всякой точки  $\mathbf{a} \in G$

$$\|\alpha - \mathbf{a}\| \geq \|\beta_G(\alpha) - \mathbf{a}\|. \quad (18)$$

<sup>2)</sup> Впрочем, до появления работы Стейна [9] этот факт считался бы совершенно очевидным.

Другим способом задания априорной информации о параметре  $\mathbf{a}$  является указание статистического веса  $p(\mathbf{a})$  каждого значения параметра, т.е. плотности априорного распределения вероятностей неизвестного значения (см. [17, 12]). В этой байесовской постановке задачи оценивания качество оценки измеряется средним риском — усреднением риска с весом  $p(\mathbf{a})$ :

$$Q_p(\Pi) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} R_{\Pi}(\mathbf{a}) p(\mathbf{a}) da_1 \dots da_s. \quad (19)$$

При байесовском подходе оптимальная оценка всегда существует и единственна (с точностью до значений на множестве исходов нулевого объема). Для ее построения известны эффективные формулы (см., например, [17], формулу (9)), а ее характеристика  $Q_p$  находится квадратурой.

Обе меры  $Q_F$  и  $Q_p$  рискованности решающего правила задаются специальным видом нормами функции риска и удовлетворяют нашим требованиям монотонности (5), калиброванности (6) и выпуклости. Однако они не инвариантны в той мере, в какой не инвариантна дополнительная информация. Как правило, в статистической литературе качество решающих процедур измеряется только описанными двумя способами [12] (см., однако, [18]). Мы последуем за традицией [19] и изучим неопределенность оценок и трудность задачи оценивания только для указанных двух постановок.

**ТЕОРЕМА 3.** Пусть  $\Pi_N$  — решающее правило для оценки  $\alpha$  неизвестного параметра  $\mathbf{a}$  семейства  $\mathfrak{R}$  по  $N$  независимым наблюдениям, и  $R_{\Pi}(\mathbf{a})$  — соответствующий риск при гауссовской функции потерь. Пусть, далее, дополнительно указан статистический вес  $p(\mathbf{a})$  априори возможных значений  $\mathbf{a}$ , интегрируемый по Риману, и качество правила предписано определять по формуле (18).

Тогда

$$Q_p(\Pi) \geq \frac{s}{N} (1 - \rho_p(N)),$$

где поправочный член

$$\rho(N) = o(1)$$

определяется только плотностью  $p(\mathbf{a})$ .

**Доказательство.** Если функция  $p(\mathbf{a})$  интегрируема по Риману, то найдется такой большой куб  $C_h$  со стороной  $2h(\varepsilon)$ , что интеграл от  $p(\mathbf{a})$  по его внешности меньше  $\varepsilon$ . Далее, найдется такое разбиение куба  $C_h$  параллельными плоскостями на достаточно малые кубики  $C_r^{(m)}$  со стороной  $2r(\varepsilon)$ , что колебание плотности  $p(\mathbf{a})$  по любому из кубиков (за исключением небольшого числа их, суммарный интеграл по которым меньше  $\varepsilon$ ) будет меньше  $\varepsilon(2h)^{-s}$ . Пусть ступенчатая функция  $p_{\varepsilon}(\mathbf{a})$  внутри каждого кубика  $C_r^{(m)}$  определена как нижняя грань  $p_m$  значений  $p(\mathbf{a})$  по этому кубику, и  $p_{\varepsilon}(\mathbf{a})$  равно нулю вне большого куба  $C_h$ . По построению  $p_{\varepsilon}(\mathbf{a}) \leq p(\mathbf{a})$  и

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} [p(\mathbf{a}) - p_{\varepsilon}(\mathbf{a})] da_1 \dots da_s \leq 3\varepsilon.$$

Проведем оценку:

$$\begin{aligned} Q_p(\Pi_N) \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} R_{\Pi}(\mathbf{a}) p(\mathbf{a}) da_1 \dots da_s &\geq \int_{-h}^h \dots \int_{-h}^h R_{\Pi}(\mathbf{a}) p_{\varepsilon}(\mathbf{a}) da_1 \dots da_s = \\ &= \sum_m p_m \int_{C_r^{(m)}} \dots \int R_{\Pi}(\mathbf{a}) da_1 \dots da_s \geq \\ &\geq \frac{s}{N} \left[ 1 - \frac{1}{r\sqrt{N}} \operatorname{th} r\sqrt{N} \right] \sum_m \int_{C_r^{(m)}} \dots \int p_{\varepsilon}(\mathbf{a}) da_1 \dots da_s \geq \frac{s}{N} \left[ 1 - \frac{1}{r\sqrt{N}} \operatorname{th} r\sqrt{N} \right] (1 - 3\varepsilon). \end{aligned}$$

При  $N > \varepsilon^{-2} r^{-2}$  множитель в квадратных скобках меньше  $1 - \varepsilon$  и, следовательно,

$$Q_p(\Pi_N) \geq \frac{s}{N} (1 - 4\varepsilon).$$

Теорема доказана (ср. [2]).

Значительно проще, простой ссылкой на теорему 1, доказывается аналогичное неравенство

$$Q_F(\Pi) \geq \frac{s}{N} (1 - \rho_F(N))$$

при измерении неопределенности по способу (17). Оно справедливо, если множество  $F$  априори возможных значений параметра  $\mathbf{a}$  содержит внутренние точки. При этом поправочный член  $\rho_F(N) \leq \frac{1}{r\sqrt{N}} \operatorname{th} r\sqrt{N}$ , где  $2r$  — сторона максимального вписанного в  $F$  куба.

Функция риска оценки  $\eta^{(N)}$  среднего арифметического наблюдений согласно (4) тождественно равна константе  $s/N$ . В силу соглашения (6) ее неопределенность не должна зависеть от способа измерения и равна также  $s/N$ . Неопределенность оптимального<sup>3)</sup> способа оценивания может быть только меньше. Поэтому справедливо неравенство

$$s(1 - \rho(N)) \leq N \inf_{\Pi} Q(\Pi_N) \leq s,$$

где  $\rho(N) \rightarrow 0$  при  $N \rightarrow \infty$ , так что асимптотическая трудность (16) рассмотренных задач действительно равна их размерности.

## § 7. Два возможных подхода к задаче точечной оценки

Здесь необходимо сделать важную оговорку. Пусть нам априори дано, что (с вероятностью единица) неизвестное значение параметра принадлежит множеству  $F$ . Что делать, если использованное наблюдателем решающее правило (например, среднее арифметическое  $\eta^{(N)}$ , или оптимальная байесовская оценка) приводит к значению, не принадлежащему  $F$ ? Возможны две различных точки

<sup>3)</sup> Для нас несущественно, достигается ли нижняя грань  $\inf Q(\Pi_N)$  или нет. Впрочем, при наших способах измерения неопределенности, существование оптимальной оценки вытекает из принципа Хелли.

зрения. При одной так же, как в теории проверки гипотез, сами такие решающие правила объявляются бессмысленными. Если априорное множество  $F$  выпукло, такая точка зрения приемлема и для нас. Мы просто заменим оценку ее “проекцией” на  $F$ . Качество оценки от этого только улучшится, а усовершенствованное правило будет уже осмысленным. Дело становится значительно сложнее, когда множество  $F$  невыпукло. В этом случае детерминированные решающие правила уже не образуют полного подкласса, лемма 1 теряет силу, так как осреднение рандомизированной оценки может оказаться бессмысленным. Оценка  $\eta^{(N)}$  также оказывается бессмысленной, и оптимальное решающее правило может приводить к неопределенности, большей  $s/N$ . Простейший пример при  $s = 1$ ,  $N = 1$  дает решетка

$$F = \{\mathbf{a} : \mathbf{a} = 0, \pm 2, \pm 4, \dots\}$$

четных целых чисел. Нетрудно проверить, что здесь

$$\sup_{\mathbf{a} \in F} R_{\Pi}(\mathbf{a}) \geq 4[1 - 2\Phi(1)] \approx 1,27 > 1.$$

Поэтому нам удобнее встать на другую точку зрения, отличную от традиционной байесовской. Наше дело — дать по наблюдениям оценку  $\alpha$ . Штраф, который нам придется впоследствии уплатить, равен  $\|\alpha - \mathbf{a}\|^2$ , какое бы  $\alpha$  из всего пространства параметров мы ни назвали. Дополнительная информация помогает нам лишь более правильно предугадать потери.

В частности, так можно поступать и когда размерность априорного множества  $F$  допустимых значений искомого параметра меньше  $s$ . Если это множество есть достаточно гладкая поверхность (с ограниченными радиусами главных кривизн), то асимптотически оптимальной оценкой параметра будет оценка наибольшего правдоподобия, получающаяся “проектированием” среднего арифметического  $\eta^{(N)}$  на  $F$ . Однако при малых  $N$  и выпуклой поверхности  $F$  несколько более выгодными могут оказаться оценки, смещенные с  $F$  в сторону центров кривизны. В обеих постановках асимптотическая трудность задачи оценивания равняется размерности поверхности  $F$ . Доказательство этого факта требует существенного усиления лемм 2 и 4, и мы его опускаем.

Заметим, наконец, что при “негладких” множествах  $F$  предел (16) может не существовать вовсе, а произведение  $N \inf Q(\Pi_N)$  будет колебаться от 0 до  $s$ . Простейшие примеры такого рода дают множества типа Канторова.

## § 8. Оптимальные решающие правила при произвольной инвариантной функции потерь

Довольно очевидно, что при совершенно произвольной функции потерь простых асимптотических закономерностей для оптимальных оценок не существует, и что простота наших формулировок во многом определяется инвариантностью гауссовской функции потерь. Оказывается, что сходные закономерности можно установить при достаточно произвольной, но инвариантной функции потерь.

Естественным обобщением гауссовской функции потерь  $\|\alpha - \mathbf{a}\|^2$  являются функции  $L(\alpha, \mathbf{a})$ , которые:

1) инвариантны, т.е.

$$L(\alpha, \mathbf{a}) = L(g\alpha, g\mathbf{a}) = L(\alpha - \mathbf{a}, \mathbf{0})$$

для любого евклидова движения  $g$  пространства параметров;

2) монотонны, т.е.

$$L(\theta\mathbf{b}, \mathbf{0}) \leq L(\mathbf{b}, \mathbf{0}) \quad \text{при } 0 \leq \theta \leq 1;$$

3) гладки, точнее, дважды дифференцируемы при  $\alpha = \mathbf{a}$ .

Соответствующую функцию риска мы будем обозначать

$$R_{\Pi}^L(\mathbf{a}) = \mathbf{M}_{\mathbf{a}}L(\alpha, \mathbf{a}),$$

а рискованность способа  $Q_F^L(\Pi)$  или  $Q_p^L(\Pi)$  вычислять, в зависимости от подхода, по формуле (17) или по формуле (19).

ЛЕММА 7. *Всякая гладкая инвариантная функция потерь*

$$L(\alpha, \mathbf{a}) = L(\alpha - \mathbf{a}, \mathbf{0}) = l(\|\alpha - \mathbf{a}\|)$$

допускает разложение

$$L(\alpha, \mathbf{a}) = c_0[L] + c[L] \cdot \sum_{j=1}^s (\alpha_j - a_j)^2 + o(\|\alpha - \mathbf{a}\|^2), \quad (20)$$

где  $c_0[L] = L(\mathbf{0}, \mathbf{0})$ , и если функция потерь  $L(\alpha, \mathbf{a})$  монотонна,  $c[L] \geq 0$ .

Без ограничения общности можно считать, что у функции потерь:

4)  $L(\mathbf{0}, \mathbf{0}) = c_0[L] = 0$ .

Если отбросить возможность  $c[L] = 0$ , то мы получим, что всякая “естественная” для семейства  $\mathfrak{A}$  функция потерь  $L(\alpha, \mathbf{a})$  локально пропорциональна гауссовскому квадрату отклонения. В частности, для любого  $\varepsilon > 0$  в достаточно малой окрестности

$$L(\mathbf{b}, \mathbf{a}) \geq c[L] \|\mathbf{b} - \mathbf{a}\|^2 (1 - \varepsilon). \quad (21)$$

Поэтому во всяком достаточно малом кубе  $C_r$  средний риск, отвечающий функции потерь  $L(\alpha, \mathbf{a})$ , не меньше

$$c[L] \frac{s}{N} \left(1 - \frac{1}{r\sqrt{N}} \text{th } r\sqrt{N}\right) (1 - \varepsilon). \quad (22)$$

В самом деле, “проекция”  $\beta_C(\alpha)$  оценки  $\alpha$  на куб  $C_r$  в силу монотонности  $L = l(\|\mathbf{b} - \mathbf{a}\|)$  приводит при  $\mathbf{a} \in C_r$  к меньшим потерям, а следовательно, и к меньшему среднему риску. Для улучшенной оценки  $\beta$  неравенство (21) справедливо, что позволяет сравнить риск  $R_{\beta}^L(\mathbf{a})$  с риском  $R_{\beta}(\mathbf{a})$ . В свою очередь неравенство для среднего риска позволяет дать нижнюю границу для рискованности решающего правила

$$Q^L(\Pi_N) \geq c[L] \frac{s}{N} (1 - \rho^L(N)),$$

где  $\rho^L(N) = o(1)$  зависит только от функции  $L$  и способа определения качества правила, т.е. либо от  $F$ , либо соответственно от  $p(\mathbf{a})$ .

Ограничение (22) определяется лишь локальным строением функции потерь. Поэтому в принципе рискованность оптимальной оценки может оказаться значительно больше. При очень быстро растущей функции потерь измерение качества оценки по риску вообще становится бессодержательным. Если множество  $F$  априори возможных законов неограничено, то всегда можно подобрать такой рост функции потерь, чтобы максимум риска и соответственно средний риск у всех оценок был бы бесконечным. (При ограниченном  $F$  можно рассматривать только оценки, спроектированные на выпуклую оболочку  $F$ , и, тем самым, иметь дело только с ограниченными потерями.) Предположим, что инвариантная функция потерь  $L(\alpha, \mathbf{a}) = l(\|\alpha - \mathbf{a}\|)$  растет не быстрее  $\exp\{k\|\alpha - \mathbf{a}\|^2\}$ . Тогда при  $N > 2k$  риск оценки наибольшего правдоподобия конечен, влияние на него больших ошибок быстро убывает и с точностью до малых высшего порядка риск равен

$$c[L] \frac{s}{N}. \quad (23)$$

Таким образом, при всякой гладкой монотонной инвариантной функции потерь, если только она не слишком быстро растет, рискованность оптимальной оценки параметра  $\alpha$  асимптотически равна (23), где  $s$  — размерность задачи,  $N$  — число наблюдений, а  $c[L]$  — константа, описывающая локальное строение функции  $L$ . При этом среднее арифметическое  $\eta^{(N)}$  является асимптотически оптимальной оценкой.

## § 9. Статистическое оценивание неизвестного закона распределения

Нам остается лишь пояснить, каким образом развитую теорию в буквально тех же формулировках удастся перенести на случай произвольных конечномерных достаточно гладких семейств законов распределения  $\mathbf{P}_a\{d\omega\}$  (см. наш препринт [6]). Соответствующим и притом несимметричным обобщением естественной гауссовской функции потерь  $\|\alpha - \mathbf{a}\|^2$  является удвоенное информационное расхождение  $L^0(\alpha, \mathbf{a}) = 2I[\mathbf{P}_a|\mathbf{P}_\alpha]$ :

$$I[\mathbf{P}'|\mathbf{P}''] = \int \left[ \ln \frac{d\mathbf{P}''}{d\mathbf{P}'}(\omega) \right] \mathbf{P}''\{d\omega\} \quad (24)$$

(см. [20], а также [21, 6]). Среднему арифметическому  $\eta^{(N)}$  наблюдений (см. (2)) соответствует оценка наибольшего правдоподобия (которую можно, правда, с натяжкой интерпретировать как оценку наименьшего информационного расхождения (см. [6])). При независимых наблюдениях оценка наибольшего правдоподобия будет асимптотически оптимальной, и ее качество опять-таки характеризуется отношением  $s/N$  размерности семейства к числу наблюдений (см. [22], а также [7]). Для важного класса геодезических экспоненциальных семейств распределений (см. [23, 12]), задаваемых на измеримом пространстве  $(\Omega, S)$  семейством плотностей

$$\frac{d\mathbf{P}_a}{d\mu}(\omega) = p(\omega; \mathbf{a}) = p_0(\omega) \exp [a^j q_j(\omega) - \Psi(\mathbf{a})], \quad (25)$$

удается провести доказательство соответствующих предложений точно по такому же плану, как для семейства  $\mathfrak{R}$ . В общем случае доказательство усложняется. Дело в том, что при нашем геометрическом подходе более естественно считать, что дается оценка  $\mathbf{P}^*$  самого неизвестного закона, а не его параметра  $\mathbf{a}$ , как обычно (см. [12, 24]). При выбранной нами функции потерь (см. (24)), формально записываемой

$$L^0 = 2I[\mathbf{P}|\mathbf{P}^*],$$

это соображение напрашивается само собой. Для “плоской поверхности” (25) всякая оценка  $\mathbf{P}^*$  при “проектировании” на поверхность только улучшается (см. [25, 6]), так что тут задача оценки закона сводится к задаче оценки параметра. В общем случае (ср. наш § 7) это не так.

Как показано в наших работах [6, 26], гладкие монотонные и инвариантные в категории FAM функции потерь  $L(\mathbf{P}_a|\mathbf{P}_\alpha)$  локально при  $\alpha = \mathbf{a} + \delta\mathbf{a}$  пропорциональны фишеровской квадратичной форме

$$L \approx c[L] \sum_{j,k} \mathbf{M}_a \left[ \frac{\partial \ln p(\omega; \mathbf{a})}{\partial a^j} \cdot \frac{\partial \ln p(\omega; \mathbf{a})}{\partial a^k} \right] \delta a^j \delta a^k, \quad (26)$$

(ср. [15, 12]). Основная идея, положенная в основу соответствующих построений, очень наглядна. Вычислим почленным интегрированием математическое ожидание приближенной (ср. [15]) функции потерь (26). По неравенству информации вторые моменты  $\mathbf{M}_a(\alpha_j - a_j)(\alpha_k - a_k)$  уклонений  $\delta\mathbf{a} = \alpha - \mathbf{a}$  оцениваются снизу через матрицу, обратную к фишеровской матрице, задающей форму (26). “Диагонализируя” эту форму (в фиксированной точке  $\mathbf{a}$ ) выбором подходящего базиса в пространстве параметров, убеждаемся, что указанная величина не менее  $c[L]s/N$ . Все остальные детали рассуждений носят технический характер, и призваны обосновать применимость хотя бы “в среднем” неравенства информации к смещенным оценкам.

Вопросам, затронутым в статье, посвящена обширная литература. Нами были процитированы только непосредственно использованные источники. Подробную библиографию с приложением исторических обзоров можно найти: по постановке задачи оценивания — в [2, 12, 27], по неравенству информации — в [7, 8, 2], по допустимым оценкам — в [10], по свойствам информационного расхождения — в [21, 6], по статистической эквивалентности и инвариантности — в [5, 6].

Наша постановка задачи статистического оценивания несколько отличается от общепринятой (ср., например, [2, 28, 12]). Мы интересуемся только асимптотическими характеристиками качества оценок, совершенно игнорируя такие полезные в практическом применении свойства, как допустимость, и делая ряд других упрощений. Зато наши формулировки ввиду их простоты очень удобны для первой грубой ориентировки.

Настоящая работа была доложена в октябре 1966 г. на семинаре по математической статистике ЛОМИ, руководимом Ю.В. Линником. Автор пользуется случаем поблагодарить Юрия Владимировича Линника за постоянное внимание. Автор благодарен также профессору Чарльзу М. Стейну за полезные беседы.

## СПИСОК ЛИТЕРАТУРЫ

1. *Gauss C.F.* Abhandlungen zur methode der kleinsten Quadrate von Carl Friedrick Gauss. — Berlin, 1887.
2. *Le Cam L.* On some asymptotic properties of maximum likelihood estimates and related Bayes' estimates // Univ. Calif. Public. Statist. 1953. V. 1, № 11. P. 277–329.
3. *Blackwell D.* Comparison of experiments, Proc. Second Berkeley Sympos // Math. Statist, and Prob. 1951. P. 93–102.
4. *Ченцов Н.Н.* Категории математической статистики // Докл. АН СССР. 1965. Т. 164, № 3. С. 511–514.
5. *Morse N., Sacksteder R.* Statistical isomorphism // Ann. Math. Statist. 1966. V. 37, № 1. P. 203–214.
6. *Ченцов Н.Н.* Инфинитезимальные методы математической статистики. — М., 1966. — (Препр. / ВИНТИ).
7. *Cramér H.* Mathematical Methods of Statistics. — Princeton University Press, 1946.
8. *Van der Varden B.L.* Mathematische Statistik. — Berlin: Springer-Verlag, 1957.
9. *Stein C.* Inadmissibility of the usual estimator for the mean of multivariate normal distribution // Proc. Third Berkeley Sympos. Math. Statist. and Prob. I. 1956. P. 197–206.
10. *James W., Stein C.* Estimation with quadratic loss function // Proc. IV Berkeley Sympos. Math. Statist. and Prob. I. 1961. P. 361–379.
11. *Stein C.* Confidence Sets for the Mean of a Multivariate Normal Distribution // J. Roy. Statist. Soc. B 24. 1962. 2. P. 265–296.
12. *Lehmann E.L.* Testing statistical hypotheses. — N.Y.: Wiley, 1960.
13. *Большев Л.Н.* Уточнение неравенства Крамера–Рао // Теория вероятн. и ее примен. 1961. Т. 6, вып. 3. С. 319–326.
14. *Hodge Jr. J.L., Lehmann E.L.* Some application of the Cramér–Rao inequality // Proc. Second Berkeley Sympos. Math. Statist. and Prob. 1951. P. 13–22.
15. *Girshick M.A., Savage L.J.* Bayes and minimax estimates for quadratic loss function // Proc. Second. Berkeley Sympos. Math. Statist. and Prob. 1951. P. 53–73.
16. *Blyth C.* On minimax statistical decision procedures and their admissibility // Ann. Math. Statist. 1951. V. 22, № 1. P. 22–42.
17. *Robbins H.* An empirical Bayes approach to statistics // Proc. Third Berkeley Sympos. Math. Statist. and Prob. I. 1956. P. 157–163.
18. *Savage L.J.* The foundation of statistics. — N.Y.: Wiley, 1954.
19. *Wald A.* Contributions to the theory of statistical estimation and testing hypotheses // Ann. Math. Statist. 1939. V. 10, № 4. P. 299–326.
20. *Kullback S., Libler R.A.* On information and sufficiency // Ann. Math. Statist. 1951. V. 22, № 1. P. 79–86.
21. *Kullback S.* Information theory and statistics. — N.Y.: Wiley, 1959.
22. *Линник Ю.В., Митрофанова Н.М.* К асимптотике распределения максимального правдоподобия // Докл. АН СССР. 1963. Т. 149, № 3. С. 518–520.
23. *Ченцов Н.Н.* К систематической теории экспоненциальных семейств распределений // Теория вероятн. и ее примен. 1966. Т. 11, вып. 3. С. 483–494.
24. *Rao M.M.* Theory of lower bounds for risk functions in estimations // Math. Ann. 1961. V. 143, № 5. P. 379–398.
25. *Kullback S., Khairat M.A.* A note on minimum discrimination information // Ann. Math. Statist. 1966. V. 37, № 1. P. 279.
26. *Ченцов Н.Н.* Инвариантные функции потерь в задачах математической статистики // Успехи матем. наук. 1967. Т. 22, вып. 1. С. 178–180.
27. *Neuman J.* Two breakthroughs in the theory of statistical decision making // Review Inst. Intern. Statist. 1962. V. 30, № 1. P. 11–27.
28. *Weiss L., Wolfowitz J.* Generalized maximum likelihood estimators // Теория вероятн. и ее примен. 1966. Т. 11, вып. 1. С. 68–93.

## О КОРРЕКТНОСТИ ЗАДАЧИ СТАТИСТИЧЕСКОГО ТОЧЕЧНОГО ОЦЕНИВАНИЯ<sup>\*)</sup>

Рассматриваемая задача статистической точечной оценки является в известном смысле обратной задачей теории вероятностей. Наблюдается случайное явление, математически описываемое некоторой вероятностной мерой  $P$  на измеримом пространстве  $(\Omega, \mathcal{J})$  всех его элементарных исходов  $\omega \in \Omega$ . Качественное описание явления: пространство  $(\Omega, \mathcal{J})$  наблюдателю известно, а мера  $P$  не известна. Ее надо оценить по наблюдениям явления.

Хорошо известно, что нетривиальные обратные задачи математической физики оказываются некорректными. Нужны дополнительные ограничения (например, требование гладкости ответа), чтобы задача получила строгое математическое содержание. В математической статистике необходимость такого ограничения более или менее узким семейством  $\mathcal{P}$  была интуитивно понята на первых же этапах ее развития. Мы обсудим здесь возникающие постановки задач на формальном уровне.

1. В математической статистике обычно имеют дело с измеримыми пространствами  $(\Omega, \mathcal{J})$  следующих трех типов:

А) когда  $\mathcal{J}$  — конечная алгебра,  $\mathcal{J} = \mathcal{J}_m$  (в этом случае  $\Omega$  является конечным объединением дизъюнктивных  $\mathcal{J}$ -атомов  $A_j$ );

В) когда  $\Omega$  является счетным объединением дизъюнктивных  $\mathcal{J}$ -атомов  $A_j$ ,  $\mathcal{J} = \mathcal{J}_\infty$ ;

С) когда  $(\Omega, \mathcal{J})$  является лебеговым измеримым пространством, т.е. взаимно однозначным и взаимно измеримым образом  $(G, \mathcal{L}_G)$ , где  $\mathcal{L}$  есть  $\sigma$ -алгебра всех абсолютно лебеговых подмножеств единичного отрезка  $E$  вещественной оси,  $G \in \mathcal{L}$ ,  $\mathcal{L}_G = \mathcal{L} \cap G$ .

В случаях В) и С) алгебры  $\mathcal{J}$  бесконечны, но порождаются последовательностями конечных подалгебр  $\mathcal{J}_m$ . Во втором случае  $\mathcal{J}_m$  имеет атомами  $A_1, \dots, A_{m-1}, A_m \cup A_{m+1} \cup \dots$ . В третьем случае за атомы  $A_{km}$  можно принять образы множеств  $G \cap \Delta_{km}$ , где  $\Delta_{km} = \{x : (k-1)m^{-1} \leq x \leq km^{-1}\}$ . Несчетно порожденными алгебрами измеримых множеств обладают континуальномерные пространства элементарных исходов, но мы их не будем касаться.

Отсутствие априорной количественной информации о  $P$  означает, что за семейство  $\mathcal{P}$  априори допустимых ответов мы должны принять совокупность всех распределений вероятностей на  $(\Omega, \mathcal{J})$ . Будем ее коротко обозначать  $\text{Cap}(\Omega, \mathcal{J})$  (по аббревиатуре от Collection of all probability measures), а через  $\text{Capd}(\Omega, \mathcal{J}, \mathcal{L})$  — подсовокупность всех тех вероятностных мер, которые обращаются в нуль на множествах из булева  $\sigma$ -идеала  $\mathcal{L} \subset \mathcal{J}$ . Когда  $\mathcal{L}$  есть

<sup>\*)</sup> Теория вероятн. и ее примен. 1981. Т. 26, вып. 1. С. 15–31. Развернутое изложение первой части доклада автора “О статистических оценках неизвестного закона распределения вероятностей” на семинаре по математической статистике, посвященном памяти члена-корреспондента АН СССР Л.Н. Большева (см.[1]).

идеал всех подмножеств  $\mu$ -меры нуль,  $\text{Card}(\Omega, \mathcal{J}, \mathcal{L})$  — подсовокупность всех  $\mu$ -доминированных вероятностных мер.

Нам будет удобно говорить, что совокупность  $\text{Car}(\Omega, \mathcal{J})$  имеет тип А) (соответственно В)), когда пространство  $\Omega$  представляется в виде конечного (соответственно счетного) объединения  $\mathcal{J}$ -атомов, т.е.  $\mathcal{J} = \mathcal{J}_m$  (или  $\mathcal{J}_\infty$ ). Совокупности  $\text{Car}(E, \mathcal{L})$  и ей изоморфным мы присвоим тип С). Всякая подсовокупность  $\text{Card}(\Omega, \mathcal{J}, \mathcal{L})$  совокупности типа В) изоморфна совокупности типа В) (или А)). Нетривиальной подсовокупности  $\text{Card}(E, \mathcal{L}, \mathcal{L}_0)$ , где  $\mathcal{L}_0$  есть идеал подмножеств лебеговой длины нуль, и всем ей изоморфным мы присвоим тип D). На более сложных алгебрах  $\mathcal{J}$ , содержащих непрерывную и дискретную части, возникают совокупности смешанного типа, т.е. смеси совокупностей чистых типов.

В рассматриваемых случаях А, В и С совокупность  $\text{Car}(\Omega, \mathcal{J})$  можно снабдить  $\sigma$ -алгеброй  $\mathcal{K}(\mathcal{J})$  колмогоровских семейств распределений вероятностей, при которой измеримое пространство  $(\text{Car}(\Omega, \mathcal{J}), \mathcal{K}(\mathcal{J}))$  также будет лебеговым типа С); см. [2, § 29]<sup>1)</sup>. В частном случае  $(\Omega, \mathcal{J})$  типа А) алгебра  $\mathcal{K}(\mathcal{J})$  является алгеброй всех абсолютно лебеговых подмножеств конечномерного симплекса  $\text{Car}(\Omega, \mathcal{J})$ . Это мезюрабеллизация совокупности вероятностных мер делает возможным вводить решающие правила в нашей задаче привычным для математической статистики образом.

**2.** Нас интересует асимптотическая “оцениваемость” закона  $P$ , когда число  $N$  наблюдаемых исходов  $\omega_i^* \in \Omega$  неограниченно растет, а доставляемая ими суммарная информация о  $P$  увеличивается. Большей частью в теоретических исследованиях наблюдения  $\omega_i^*$  считают статистически независимыми. Иногда их предполагают стационарно связанными с одним и тем же маргинальным распределением  $P$ . Минимальное ограничение, при котором еще можно ожидать состоятельной оценки  $P$ , — выполнение закона больших чисел для атомарных событий  $A_{jm}$  из порождающих алгебр  $\mathcal{J}_m$ . Мы наложим несколько более сильное ограничение в виде

$$P^{(N)} \left\{ \sum_{j=1}^m \left| \frac{\nu(A_{jm}|N)}{N} - P\{A_{jm}\} \right| > \varepsilon \right\} \leq C \frac{m^\gamma}{\varepsilon^\alpha N^\beta}, \quad (1)$$

где  $P^{(N)}$  — совместное вероятностное распределение на  $(\Omega^N, \mathcal{J}^N)$  первых  $N$  наблюдений  $\omega_1^*, \dots, \omega_N^*$ ,  $\nu(D|N)$  — число осуществлений событий  $D$  в указанных наблюдениях. Связь  $P^{(N)}$  (совместного распределения наблюдений, отвечающего оцениваемому закону  $P$ ) и  $P$  при этом не фиксируется. В частности, не требуется, чтобы маргинальные распределения точно совпадали с  $P$ ; (1) означает лишь совпадение “в среднем”.

Зато предполагается, что положительные показатели  $\alpha, \beta, \gamma$  (и, быть может, константа  $C$ ) известны (конструктору решающего правила). В частности, для схемы независимых испытаний

$$\alpha = 2, \quad \beta = 1, \quad \gamma = 1, \quad C = 1; \quad (2)$$

<sup>1)</sup>Эта мезюрабеллизация  $\text{Car}(\Omega, \mathcal{J})$  является развитием подходов [3] и [4].

это следует из неравенства Чебышева и оценки

$$\mathbf{M} \left[ \sum_j \left| N^{-1} \nu(A_{jm}|N) - P\{A_{jm}\} \right| \right]^2 \leq mN^{-1}.$$

Обычно при зависимых испытаниях имеет место более медленная сходимость (с  $\beta \leq 1$ ). Однако при квазислучайном моделировании случайных явлений, используемом в методе Монте-Карло, удается добиться в (1) и несколько лучшей, чем с (2), сходимости частот.

Для совокупностей типа А) и В) множество кортежей  $(\omega_1^*, \dots, \omega_N^*)$  исходов не более чем счетно, а мощность совокупности распределений — континуум. Поэтому никакой конечный эксперимент (без дополнительной априорной информации) не может точно указать  $P$ . Иное дело совокупность  $\text{Cap}(E, \mathcal{L})$ . Она также имеет мощность континуума и может быть взаимно однозначно и взаимно измеримо отображена  $P \rightarrow z(P)$  на некоторое измеримое подмножество  $H$  единичного отрезка (на этом, собственно, и основана конструкция колмогоровской алгебры семейств распределений вероятностей). А элементарных исходов  $x \in E$  тоже континуум. И если первое испытание распределено по закону  $P^{(1)}$ :  $P_P^{(1)}\{\xi = z(P)\} = 1$ , то первое же наблюдение доставляет полную информацию о  $P$ . В терминах асимптотики частот подобную патологическую “подсказку” исключить нельзя. Поэтому мы требуем, чтобы соответствие  $P \rightarrow P^{(N)}$  сохраняло следующие статистические связи:

а) при любых  $P \in \mathcal{P}$  и  $N$

$$P\{D\} = 0 \implies P^{(N)}\{\nu(D|N) > 0\} = 0; \quad (3)$$

б) при любых  $P, Q \in \mathcal{P}$  и образующих  $A_{j(1)}, \dots, A_{j(N)}$  алгебры  $\mathcal{J}$

$$\begin{aligned} [P\{\cdot|A_{j(k)}\} = Q\{\cdot|A_{j(k)}\}, \quad k = 1, \dots, N] \implies \\ \implies [P^{(N)}\{\cdot|A_{j(1)} \times \dots \times A_{j(N)}\} = Q^{(N)}\{\cdot|A_{j(1)} \times \dots \times A_{j(N)}\}]. \end{aligned} \quad (4)$$

Для схемы независимых испытаний условия (3) и (4) выполнены очевидным образом.

**3.** Мы будем рассматривать последовательности оценок  $P_N^*$ ,  $N = 1, 2, \dots$ , как детерминированных

$$P_N^*\{\cdot\} = P_N^*\{\cdot|\omega_1^*, \dots, \omega_N^*\}, \quad (5)$$

так и рандомизированных

$$P_N^*\{\cdot\} = P_N^*\{\cdot|\omega_1^*, \dots, \omega_N^*; \eta_N\}, \quad (6)$$

где  $\eta_N$  — случайный рандомизируемый параметр. Рандомизированные оценки  $P_N^*$  задаются (см. [2]) марковскими переходными распределениями вероятностей  $\Pi_N(\omega_1, \dots, \omega_N; d[Q\{\cdot\}])$  из  $(\Omega^N, \mathcal{J}^n)$  в  $\text{Cap}(\text{Cap}(\Omega, \mathcal{J}), \mathcal{K}(\mathcal{J}))$ . При этом

$$\forall (\omega_1, \dots, \omega_N) \in \Omega^N \quad \Pi(\omega_1, \dots, \omega_N; \text{Cap}(\Omega, \mathcal{J})) = 1.$$

Для интересующих нас простейших типов подсовокупностей  $\mathcal{P} \subset \text{Car}(\Omega, \mathcal{J})$  обычное дополнительное требование:  $P_N^* \in \mathcal{P}$  почти наверное, т.е.

$$(P^{(N)} \circ \Pi_N)\{\mathcal{P}\} = \int P^{(N)}\{d\omega_1 \times \dots \times d\omega_N\} \Pi_N(\omega_1, \dots, \omega_N; \mathcal{P}) = 1,$$

излишне. Эти  $\mathcal{P}$  выпуклы, и проектированием состоятельной оценки на  $\mathcal{P}$  мы только улучшим ее свойства.

4. За меру погрешности оценки  $P_N^*$  мы примем вариацию уклонения меры  $P_N^*\{\cdot\}$  от истинного закона  $P\{\cdot\}$ :

$$|P_N^* - P| = \sup \sum_j |P_N^*\{A_j\} - P\{A_j\}|, \quad (7)$$

где верхняя грань берется по всем конечным подалгебрам  $\mathcal{J}' \subseteq \mathcal{J}$ , а суммирование производится по всем атомам  $A_j$  алгебры  $\mathcal{J}'$ . Для лебеговых измеримых пространств  $(\Omega, \mathcal{J})$  этот выпуклый функционал  $\mathcal{K}$ -измерим и равен пределу значений сумм (7) по последовательности порождающих алгебру  $\mathcal{J}$  подалгебр  $\mathcal{J}_m$  (см. [2, 5, 6]). Вариация не только естественная норма в пространстве зарядов (т.е. мер, принимающих значения обоих знаков), задающая в нем \*-сильную топологию. Она также единственная (с точностью до множителя) норма в этом пространстве, задающая на  $\text{Car}(\Omega, \mathcal{J})$  метрику, инвариантную относительно категории всех статистических решающих правил; см. подробнее [2, 7].

Естественная характеристика непохожести двух вероятностных законов  $P$  и  $Q$  — относительная энтропия, или информационное уклонение Кульбака–Лейблера–Санова  $\mathbf{I}[Q : P]$ . Если  $P \gg Q$ , то

$$\mathbf{I}[Q : P] = \int_{\Omega} \left[ \frac{dQ}{dP}(\omega) \ln \frac{dQ}{dP}(\omega) \right] P\{d\omega\}, \quad (8)$$

и  $\mathbf{I}[Q : P] = +\infty$  в противном случае. Эта величина несимметрична, но само отношение похожести случайных явлений несимметрично, и его надо мерить несимметричным образом; см. наши объяснения в [7]. Оказывается, что  $2\mathbf{I}[Q : P]$  является несимметричным обобщением квадрата евклидова расстояния, удовлетворяющим несимметричному равенству Пифагора. Оценивание неизвестного закона  $P$  при функции потерь  $2\mathbf{I}[P_N^* : P]$  обобщает гауссов метод наименьших квадратов для семейства нормальных законов с единичной матрицей ковариаций.

Чисар рассмотрел в [3, 4] целый класс сходных уклонений. Пусть  $\varphi : \mathbb{R}^+ \rightarrow \mathbb{R}$  — выпуклая функция:

$$\alpha_i \geq 0 \quad (\forall i), \quad \sum \alpha_i = 1 \implies \sum \alpha_i \varphi(x_i) \geq \varphi\left(\sum \alpha_i x_i\right),$$

нелинейная в окрестности точки  $x = 1$ :

$$\begin{aligned} \varphi(1) = 0, \quad \exists a: \quad \forall x \neq 1 \quad \varphi(x) - a(x-1) > 0, \\ \lim_{x \rightarrow 0} \varphi(x) = b \leq +\infty, \quad \lim_{x \rightarrow \infty} \frac{\varphi(x)}{x} = c \leq +\infty. \end{aligned}$$

Положим

$$\Phi = \Phi_\varphi[Q : P] = \sup \sum_j P\{A_j\} \varphi\left(\frac{Q\{A_j\}}{P\{A_j\}}\right) \quad (9)$$

с правилом раскрытия неопределенностей, отличным от принятого в теории интеграла Лебега:

$$0 \cdot \varphi\left(\frac{0}{0}\right) = 0, \quad \forall z > 0 \quad 0 \cdot \varphi\left(\frac{z}{0}\right) = zc. \quad (10)$$

Верхняя грань в (9) берется по всем конечным подалгебрам  $\mathcal{J}' \subseteq \mathcal{J}$ , а сумма — по всем атомам подалгебры  $\mathcal{J}'$ . При правиле (10) суммы в (9) монотонны по  $\mathcal{J}'$ . В лебеговых пространствах  $(\Omega, \mathcal{J})$  монотонный предел (9) достигается на последовательности порождающих алгебр  $\mathcal{J}_m$ ,  $m = 1, 2, \dots$  (ср. [2, 5]), так что все отклонения вида (9)  $\mathcal{K}(\mathcal{J})$ -измеримы. При  $\varphi(x) = x \ln x$  получается  $\Phi = \mathbf{I}[Q : P]$ , а при  $\varphi(x) = |x - 1|$  — вариация  $\mathbf{V}[Q : P] = |Q - P|$ . Для всех  $\Phi_\varphi$  указанного класса  $\Phi[P : P] = 0$  и  $\Phi[Q : P] > 0$  при  $Q \neq P$ . Когда пределы  $b$  и  $c$  конечны, окрестности  $\{Q : \Phi[Q : P] < r\}$  порождают ту же топологию, что и вариация; в противном случае — более сильную псевдотопологию. При  $c = +\infty$

$$\Phi[Q : P] = +\infty,$$

когда  $P$  не доминирует  $Q$ , а при  $b = +\infty$ , когда  $Q$  не доминирует  $P$ .

5. Обобщая общепринятую терминологию [8], условимся называть последовательность решающих правил  $\Pi_N$  построения оценок  $P_N^*$  неизвестного закона  $P$  из априорного семейства  $\mathcal{P}$  по  $N$  наблюдениям,  $N = 1, 2, \dots$ , *сильно состоятельной по вероятности*, если

$$\lim_{N \rightarrow \infty} (P^{(N)} \circ \Pi_N)\{|P_N^* - P| < \varepsilon\} = 1 \quad \forall \varepsilon > 0, \quad \forall P \in \mathcal{P}, \quad (11)$$

и *сильно состоятельной почти наверное*, если

$$(P^{(\infty)} \circ \Pi)\{|P_N^* - P| \rightarrow 0\} = 1 \quad \forall P \in \mathcal{P}. \quad (12)$$

**ОПРЕДЕЛЕНИЕ 1.** Назовем задачу статистической оценки неизвестного вероятностного закона из априорного семейства  $\mathcal{P}$  по последовательности наблюдений *сильно корректной по вероятности* (соответственно *почти наверное*), если она обладает хотя бы одной сильно состоятельной по вероятности (соответственно почти наверное) последовательностью решающих правил.

Для практических потребностей интересны в первую очередь только такие процедуры, о которых известно заранее, сколько наблюдений надо обработать, чтобы при заданном уровне доверия гарантируемая погрешность ответа не превышала  $\varepsilon$ . А это подразумевает равномерную состоятельность решающего правила.

**ОПРЕДЕЛЕНИЕ 2.** Назовем задачу статистической оценки для априорного семейства  $\mathcal{P}$  *равномерно сильно корректной*, если она обладает последовательностью решающих правил  $\Pi_N$ , равномерно состоятельных на  $\mathcal{P}$ , т.е.

$$\lim_{N \rightarrow \infty} \inf_{P \in \mathcal{P}} (P^N \circ \Pi_N)\{|P_N^* - P| < \varepsilon\} = 1 \quad \forall \varepsilon > 0. \quad (13)$$

Заметим, что существование равномерно состоятельной по вероятности последовательности  $\Pi_N$  влечет равномерную состоятельность почти наверное некоторой ее подпоследовательности, так что оба эти вида равномерной корректности совпадают.

Удобной характеристикой точности решающего правила  $\Pi_N$  является риск (см. [9]). Принимая, что потери от погрешности оценки  $P_N^*$  определяются функционалом  $\Phi[P_N^* : P]$ , мы приходим к риску

$$\mathfrak{R}_{\Pi}(P) = \int_{\Omega^N} \int_{\text{Car}(\Omega, \mathcal{J})} \Phi[Q : P] \Pi(\omega_1, \dots, \omega_N; d[Q\{\cdot\}]) P^{(N)} \left\{ \bigotimes_{i=1}^N d\omega_i \right\}. \quad (14)$$

Заметим, что  $\mathbf{V}[P_N^* : P] = |P_N^* - P|$  обобщает лапласов модуль погрешности, а  $2\mathbf{I}[P_N^* : P]$  — гауссову сумму квадратов погрешностей (см. [10, 7]).

Условимся говорить, что последовательность решающих правил  $\Pi_N = \Pi(N)$  приводит к асимптотически нулевому риску или просто  $\Phi$ -ультрасостоятельна на  $\mathcal{P}$ , если

$$\mathfrak{R}_{\Pi(N)}(P) \rightarrow 0, \quad \forall P \in \mathcal{P}.$$

**ОПРЕДЕЛЕНИЕ 3.** Назовем задачу статистической оценки неизвестного вероятностного закона  $P$  из априорного семейства  $\mathcal{P}$  по последовательности наблюдений  $\Phi$ -ультрасильно корректной, если она допускает  $\Phi$ -ультрасостоятельную на  $\mathcal{P}$  последовательность решающих правил  $\Pi(N)$ , и равномерно  $\Phi$ -ультрасильно корректной, когда (при  $N \rightarrow \infty$ )

$$\sup_{P \in \mathcal{P}} \mathfrak{R}_{\Pi(N)}(P) \rightarrow 0.$$

Аналогичным образом можно определить слабую (и ультраслабую) состоятельность последовательности оценок относительно какой-либо метрики (см. [11, 12]), задающей слабую топологию на пространстве зарядов. Например, известные теоремы В.И. Гливенко [13] и А.Н. Колмогорова [14] устанавливают слабую состоятельность почти наверное эмпирической меры  $F_n^*\{\cdot\}$  как оценки теоретической меры при  $\mathcal{P} = \text{Car}(E, \mathcal{L})$  и ее равномерную слабую состоятельность при  $\mathcal{P} = \text{Car}(E, \mathcal{L}, \mathcal{L}_0)$ . Однако наша цель — рассмотреть именно сильную и ультрасильную корректность задачи оценивания.

**6. ТЕОРЕМА 1.** Пусть совместные распределения  $P^{(N)}$  вероятностей последовательности наблюдений удовлетворяют условию (1) равномерно по  $P \in \mathcal{P}$  с известными показателями  $\alpha, \beta, \gamma$ .

Тогда задача статистической оценки  $P \in \mathcal{P}$  по наблюдениям:

- 1) равномерно сильно корректна для  $\mathcal{P} = \text{Car}(\Omega, \mathcal{J}_m)$  с конечной алгеброй  $\mathcal{J}_m$ , т.е. для любой совокупности типа А) нашей классификации;
- 2) сильно корректна для  $\mathcal{P} = \text{Car}(\Omega, \mathcal{J}_\infty)$  со счетно дискретной алгеброй  $\mathcal{J}_\infty$  (тип В));
- 3) сильно корректна для  $\mathcal{P} = \text{Car}(E, \mathcal{L}, \mathcal{L}_0)$  и ей изоморфных (тип D)).

**Доказательство.** 1) Обозначим всюду далее  $F_N^*\{\cdot\}$  эмпирическую меру, построенную по  $N$  наблюдениям:

$$F_N^*\{D\} = N^{-1}[\chi_D(\omega_1^*) + \dots + \chi(\omega_N^*)] = N^{-1}\nu(D|N). \quad (15)$$

Равномерная сильная состоятельность по вероятности оценки  $P_N^* = F_N^*$  следует из предположения (1).

2) Для  $\mathcal{J} = \mathcal{J}_\infty$  справедлива оценка

$$|F_N^* - P| \leq \sum_{j=1}^m |F_N^*\{A_{jm}\} - P\{A_{jm}\}| + 2 \sum_{k=m}^{\infty} P\{A_k\}. \quad (16)$$

Обозначим  $l = 1 + \Pi[(1 + \gamma)\beta^{-1}]$ , где  $\Pi(x)$  — целая часть  $x$ , и возьмем подпоследовательность значений  $N(m) = m^l$ . Для этой подпоследовательности сумма правых частей в (1) конечна при любом  $\varepsilon > 0$ . Поэтому по лемме Бореля–Кантелли первое слагаемое в правой части (16) стремится к нулю почти наверное. Второе (неслучайное) слагаемое в (16) стремится к нулю как остаток сходящегося ряда. Значит, оценка  $P_N^* = F_{n(N)}^*$ , где  $m^l = n(N) \leq N < (m + 1)^l$ , сильно состоятельна почти наверное.

Доказательство 3) несущественно огрубляет рассуждения Надарая (см. [15]). Пусть  $p(x) = dP/dx$ ,  $p \in L_1$ . Возьмем  $m = 2^k$ . Построим по  $N$  наблюдениям гистограмму  $\pi_{mN}(x)$  относительно разбиения  $E$  на отрезки  $A_{jm(k)} = [(j - 1)2^{-k}, j2^{-k}]$  и отвечающую ей меру  $P_{mN}^*$ :  $dP_{mN}^*/dx = \pi_{mN}(x)$ . Пусть, наконец,  $p_{m(k)}(x)$  — кусочно постоянная плотность,  $p_m(x) = mP\{A_{jm}\}$  при  $x \in A_{jm}$  и  $P_{m(k)}\{A_{jm}\} = P\{A_{jm}\} \forall j$ , где  $dP_m/dx = p_m(x)$ . Тогда справедлива оценка

$$|P_{mN}^* - P| \leq \sum_{j=1}^m |F_N^*\{A_{jm}\} - P\{A_{jm}\}| + |P_m - P|. \quad (17)$$

Отрезки  $A_{jm(k)} \forall j, k$ , образуют счетный базис алгебры  $\mathcal{L}/\mathcal{L}_0$ , поэтому  $|P_{m(k)} - P| \rightarrow 0$  при  $k \rightarrow \infty$  (см., например, [16]). Выберем теперь подпоследовательность значений  $N(k) = m^l = 2^{kl}$ , где  $l = 1 + \Pi[(1 + \gamma)\beta^{-1}]$ . Повторяя рассуждения, использованные в доказательстве 2), убеждаемся, что и первое слагаемое в (17) сходится к нулю почти наверное. Значит, подпоследовательность оценок  $P_{m(k), N(k)}^*$  сильно состоятельна почти наверное.

**7.** Прежде чем формулировать теорему 2, напомним понятие  $\varepsilon$ -емкости компакта и его свойства (см. [17, 18]). Подмножество  $U$  компакта  $W$ , лежащего в метрическом пространстве  $R$ , называется  $\varepsilon$ -различимым, если любые две его различные точки лежат на расстоянии, большем  $\varepsilon$ . Обозначим  $M_\varepsilon(W)$  максимально возможное число точек в  $\varepsilon$ -различимом подмножестве множества  $W$ ;  $\log_2 M_\varepsilon(W)$  называется  $\varepsilon$ -емкостью компакта  $W$ . Это название объясняется следующим образом. Окружим каждую точку  $y$   $\varepsilon$ -различимого подмножества  $U$  шаром  $O_{\varepsilon/2}(y)$  радиуса  $\varepsilon/2$  с центром в точке  $y$ . Тогда шары  $O_{\varepsilon/2}(y)$  не пересекают друг друга. Таким образом,  $M_\varepsilon(W)$  есть максимальное число непересекающихся шаров радиуса  $\varepsilon/2$  (и диаметра  $\varepsilon$ ), которые можно вложить в  $W$ . Величину  $M_\varepsilon(W)$  можно оценить снизу. Пусть  $W$  дискретно и внутрь каждого шара  $O_\varepsilon(x)$  радиуса  $\varepsilon$  с центром в  $x \in W$  попадает не более  $v_\varepsilon$  точек из  $W$ ; тогда

$$M_\varepsilon(W) \geq (v_\varepsilon)^{-1} \text{Card } W. \quad (18)$$

В самом деле, окружим каждую точку  $y$   $\varepsilon$ -различимого  $U$  шаром  $O_\varepsilon(y)$ . Если их объединение не покрывает всего множества  $W$ , то  $U$  можно увеличить. Значит,

для максимального  $U$

$$\text{Card } W \leq v_\varepsilon \text{Card } U.$$

Хорошо известен принцип Дирихле: если в  $M$  непересекающихся ящиков кладут  $K < M$  предметов, то хоть один из ящиков останется пустым. Нам потребуется его следующее менее известное обобщение: если  $K$  буханками хлеба хотят накормить  $M$  человек, где  $M > 2K$ , то хотя бы одному человеку достанется меньше полбуханки<sup>2)</sup>.

ЛЕММА 1. Пусть даны две субстохастические (в частности, стохастические) матрицы  $(p_{ij})_{i,j=1}^{i=M,j=K}$  и  $(\pi_{ji})_{j,i=1}^{j=K,i=M}$ , где  $M > 2K$ .

Тогда найдется такой индекс  $i$ , что

$$\sum_{j=1}^K \pi_{ji} < \frac{1}{2}, \quad \sum_{j=1}^K p_{ij} \pi_{ji} < \frac{1}{2}. \quad (19)$$

Более точную оценку дает неравенство чебышевского типа

$$\text{Card} \left\{ i : \sum_j \pi_{ji} < \varepsilon \right\} > M - K\varepsilon^{-1},$$

откуда немедленно следует первое из неравенств (19). Второе же вытекает из первого, так как все  $p_{ij} \leq 1$ .

8. ТЕОРЕМА 2. Задача статистической оценки  $P \in \mathcal{P}$  по наблюдениям не является равномерно сильно корректной для совокупностей:

1)  $\mathcal{P} = \text{Card}(\Omega, S_\infty)$ , т.е. для совокупностей типа В) нашей классификации, если совместные распределения вероятностей  $P_P^{(N)}\{\cdot\}$  кортежей  $(\omega_1^*, \dots, \omega_N^*)$  наблюдений удовлетворяют условию (3);

2)  $\mathcal{P} = \text{Card}(E, \mathcal{L}, \mathcal{L}_0)$  (тип D)), если совместные распределения  $P_P^{(N)}\{\cdot\}$  удовлетворяют условиям (3) и (4).

Доказательство. 1) Рассмотрим подсовокупность всех распределений вероятностей  $P$ , сопоставляющих меру 0 всем атомам  $A_k$  при  $k \geq t + 1$ . В наших обозначениях это  $\text{Card}(\Omega, \mathcal{J}_\infty, \mathcal{L}_m)$ , где  $\mathcal{L}_m = \mathcal{J} \cap A'_m$ ,  $A'_m = A_{m+1} \cup A_{m+2} \cup \dots$ . По условию (3) при  $P \in \text{Card}(\Omega, \mathcal{J}_\infty, \mathcal{L}_m)$  с  $P^{(N)}$ -вероятностью единица событие  $A'_m$  ни разу не произойдет. Следовательно,  $N$  наблюдений могут привести всего к

$$K = m^N \quad (20)$$

существенно различающимся исходам  $A_j = A_{j(1)} \times \dots \times A_{j(N)}$  эксперимента,  $j(k) \leq m$ ,  $k = 1, \dots, N$ . Здесь два элементарных исхода  $(\omega'_1, \dots, \omega'_N)$  и  $(\omega''_1, \dots, \omega''_N)$ , принадлежащие одному атому  $A_j \in \mathcal{J}^N$ , неразличимы в  $\mathcal{J}^N$ , но порядок осуществления атомарных событий может быть существенным.

Оценим теперь  $\varepsilon$ -емкость совокупности  $W = \text{Card}(\Omega, \mathcal{J}_\infty, \mathcal{L}_m)$ . Рассмотрим конечное семейство  $W' \subset W$  распределений  $P$ , у каждого из которых вероятности каких-то  $lr$  атомов равны  $1/lr$ , а остальных — равны нулю. Всего таких  $P$

<sup>2)</sup>В литературе неоднократно обсуждался частный случай  $M = 5000$ ,  $K = 5$  (см., например, Евангелие от Матфея, гл. XIV).

будет

$$\text{Card } W' = C_m^{lr} = m(m-1)\dots(m-lr+1)[(lr)!]^{-1}. \quad (21)$$

Два таких закона различаются по вариации не больше чем на  $2/r$  тогда и только тогда, когда у них совпадают не менее  $l(r-1)$  положительных атомов. Поэтому число  $v_{2/r}$  “точек” из  $W'$ , попадающих в шар радиуса  $2/r$  с центром в  $P \in W'$ , есть

$$v_{2/r} = C_{lr}^0 + C_{lr}^1 + \dots + C_{lr}^l. \quad (22)$$

Из (21) и (22) по (18) следует

$$M_{2/r}(W') = B(l, r)m^{lr} + O(m^{lr-1}), \quad m \rightarrow \infty. \quad (23)$$

Возьмем  $l = N$ ,  $r = 2$ . Тогда согласно (20) и (23)  $K/M_1 \rightarrow 0$  при  $m \rightarrow \infty$ , и мы можем выбрать такое  $m$ , что  $M_1 > 2K$ .

Покажем теперь, что задача оценки не может быть равномерно состоятельной. Точнее, мы покажем, что для любой последовательности решающих правил  $\Pi_N$  при любом числе наблюдений  $N$  найдется закон  $P \in \text{Cap}(\Omega, \mathcal{J}_\infty)$ , для которого

$$(P^{(N)} \circ \Pi_N) \left\{ |P_N^* - P| \leq \frac{1}{2} \right\} < \frac{1}{2}, \quad (24)$$

что противоречит требованию (13) определения 2. Выберем, как указано выше,  $m$  так, чтобы  $M_1 > 2K$ , построим семейство  $W'$  описанных выше законов и выберем в нем максимальное 1-различимое подсемейство  $U$ . По выбору  $m$

$$\text{Card } U = M_1 > 2K.$$

Если каждую точку  $P_i \in U$  окружить шаром  $O_i$  радиуса  $1/2$ , то эти шары не будут попарно пересекаться. Пусть правило  $\Pi_N$  детерминировано. Тогда по каждому кортежу наблюдений  $(\omega_1^*, \dots, \omega_N^*) \in A_{j(1)} \times \dots \times A_{j(N)}$  оно дает одну и ту же оценку  $P_N^* = Q_{\mathbf{j}}$ ,  $\mathbf{j} = (j(1), \dots, j(N))$ . Так как число выводов  $K < M_1$ , то хотя бы в один шар  $O_i$  не попадет ни одного  $Q_{\mathbf{j}}$ . Значит, при  $P^{(N)}$ -почти всех исходах будет  $|P_N^* - P_i| > 1/2$ , что дает (24) даже с нулем в правой части неравенства вместо  $1/2$ . Пусть правило  $\Pi_N$  рандомизировано. Построим по нему субстохастическую матрицу

$$\pi_{ji} = \Pi(O_i | A_{j(1)} \times \dots \times A_{j(N)}).$$

По лемме 1 найдется такой шар  $O_i$ , для которого

$$(P^{(N)} \circ \Pi_N) \{O_i\} = \sum_j p_{ij} \pi_{ji} < 1/2,$$

что эквивалентно (24).

2) Совокупность  $\text{Card}(E, \mathcal{L}, \mathcal{L}_0)$  содержит семейство  $W$  распределений вероятностей, конгруэнтное относительно категории статистических решающих правил только что рассмотренной совокупности  $\text{Cap}(\Omega, \mathcal{J}_\infty)$ . Например,

$$W = \left\{ P\{\cdot\} : P = \sum p_k Q_k \quad \forall k, \quad p_k \geq 0, \quad \sum p_k = 1 \right\}, \quad (25)$$

где  $dQ_k/dx = 2^k$  при  $x \in A_k = [2^{-k}, 2^{-k+1})$  и  $dQ_k/dx = 0$  вне этого полуинтервала. Алгебра  $\mathcal{J}_\infty$  с атомами  $A_k$  является по построению достаточной алгеброй семейства  $W$ . Отсюда по условию (4) алгебра  $\mathcal{J}_\infty^N$  будет достаточной алгеброй для семейства  $W^{(N)}$  законов распределения  $P^{(N)} = P_P^{(N)}\{\cdot\}$  вероятностей кортежей  $(\omega_1^*, \dots, \omega_N^*)$  наблюдений. Следовательно,  $W^{(N)}$  конгруэнтно  $[\text{Car}(\Omega, \mathcal{J}_\infty)]^N$ . Это означает, в частности, что всякое решающее правило  $\Pi_N$  для семейства  $W^{(N)}$  на  $(E^N, \mathcal{L}^N)$  переходит в решающее правило  $\Pi'_N$  для  $\text{Car}(\Omega^N, \mathcal{J}_\infty^N)$  с теми же статистическими свойствами:

$$\Pi'_N(A_j; \cdot) = \int_A \Pi(x_1, \dots, x_N; \cdot) P^{(N)}\{dx_1 \times \dots \times dx_N | A_j\},$$

где  $A = A_j = A_{j(1)} \times \dots \times A_{j(N)}$ , а  $P^{(N)}\{\cdot | A_j\}$  совпадает у всех  $P \in W$ . Так как расстояние по вариации инвариантно относительно категории статистических решающих правил, то неравенство (24) будет справедливо для некоторого  $P \in W \subset \text{Card}(E, \mathcal{L}, \mathcal{L}_0)$ .

**9.** Рассмотрим возможности оценивания по наблюдениям неизвестного закона  $P$  на  $(E, \mathcal{L})$  из более широких априорных совокупностей, чем  $\text{Card}(E, \mathcal{L}, \mathcal{L}_0)$ . Для простоты ограничимся случаем независимых наблюдений. Заметим, что при независимых испытаниях наблюдавшийся исход  $x \in E$  маргинальной вероятности нуль с вероятностью единица вторично наблюдаться не может. Если же его маргинальная вероятность  $P\{x\} > 0$ , то он будет наблюдаться многократно с частотой, асимптотически приближающейся к  $P\{x\}$ . Отсюда вытекает следующий результат.

**ТЕОРЕМА 3.** Пусть  $\mathcal{P}_0$  есть совокупность всех тех вероятностных мер на  $(E, \mathcal{L})$ , у которых есть только абсолютно непрерывная (по лебеговой длине  $\Lambda$ ) и чисто дискретная компоненты. Для  $\mathcal{P}_0$  задача статистической точечной оценки по независимым наблюдениям сильно корректна.

**Доказательство.** Расположим  $N$  полученных наблюдений в вариационный ряд, каждую группу совпавших наблюдений используем для оценки вероятности соответствующего атома, а по оставшейся цензурированной выборке построим гистограмму (нормированную множителем  $N^{-1}$ ). Доказательство сходимости оценки удобно получить, представив искомое  $P$  в виде суммы  $P = pR_1 + qR_2 + rR_3$ , где  $pR_1$  — абсолютно непрерывная компонента  $P$ ,  $qR_2$  — мера, сосредоточенная в конечном числе высоко вероятных атомов,  $rR_3$  — остаток дискретной компоненты с  $rR_3\{E\} < \varepsilon$ . Затем каждый случайный выбор  $x^*$  по закону  $P$  следует интерпретировать как выбор одного из  $R_i$  ( $i = 1, 2, 3$ ) с вероятностями  $p, q, r$  соответственно и последующий независимый розыгрыш  $x^* \in E$  по выбранному закону  $R_i\{dx\}$ .

**10.** Решающее правило теоремы 3 может быть обобщено. Пусть  $\xi$  — равномерно распределенная на  $[0, 1]$  случайная величина с двоичным разложением  $\xi = 0, \varkappa_1 \varkappa_2 \dots$ . Введем параметр  $\eta$ ,  $0 \leq \eta \leq 1$ ,  $\eta = 0, \lambda_1 \lambda_2 \dots$ . Пусть случайная величина  $\zeta = \zeta_\eta$  есть  $\zeta = \varphi(\xi) = 0, \varkappa_1 \lambda_1 \varkappa_2 \lambda_2 \dots$ . Она имеет распределение  $P_\eta = \Lambda \varphi_\eta^{-1}\{\cdot\}$ . Все  $P_\eta$  сосредоточены на  $[0, 1]$ , непрерывны и сингулярны друг другу и лебеговой мере  $\Lambda$ . Очевидно, идентифицировать  $P \in \mathcal{P} = \{P_\eta, 0 \leq \eta \leq 1\}$  можно по единственному наблюдению. Смеси же законов из семейства  $\mathcal{P}$  и семейств

ва  $\text{Card}(E, \mathcal{L}, \mathcal{L}_0)$  могут оцениваться методом, аналогичным использованному в теореме 3.

Однако вся совокупность  $\text{Card}(E, \mathcal{L})$  слишком широка, чтобы задача статистической точечной оценки была для нее корректной. Опишем сначала некоторые семейства непрерывных сингулярных мер, которые “трудно” отличить от абсолютно непрерывных.

Пусть  $N_1 < N_2 < \dots$  — любая возрастающая последовательность натуральных чисел. Построим по ней монотонную последовательность  $\{m_k\}$  целых чисел с  $m_1 = 0$  по правилу

$$m_{k+1} - m_k = n_{k+1} = 4 + C[\log_2 k(k+1) + \log_2 N_k(N_k+1)]. \quad (26)$$

Носитель сингулярной меры определяется рекуррентно. Пусть на  $k$ -м шаге его покрытие  $D_k$  состояло из  $2 \uparrow (m_k - k)$  двоично-рациональных интервалов  $\Delta_k$  вида  $[(l-1)2 \uparrow (-m_k), l2 \uparrow (-m_k)]$ ; здесь  $2 \uparrow n = 2^n$ . На  $(k+1)$ -м шаге интервал  $\Delta_k$  делится на  $2 \uparrow n_{k+1}$  равных частей, и из них произвольно отбирается половина. После такой операции над каждым интервалом, входящим в  $D_k$ , получается  $2 \uparrow (m_{k+1} - k - 1)$  двоично-рациональных интервалов, составляющих  $D_{k+1}$ . Припишем каждому выбранному интервалу  $\Delta_k$  меру  $2 \uparrow (-m_k + k)$ . Такое задание согласовано по  $k$  и определяет на  $(E, \mathcal{L})$  сингулярную непрерывную меру с носителем  $\cap D_k$ . Итак, каждой последовательности  $\{D_k\}$  такого рода вложенных покрытий отвечает единственная вероятностная мера с “равномерным” распределением на  $\cap D_k$ . В построенное семейство сингулярных мер можно ввести следующую измеримую параметризацию. Пусть  $D_k = \cup \Delta_{kj}$ ,  $2M_{k+1} = 2 \uparrow n_{k+1}$  — число интервалов, на которое делится каждый  $\Delta_{kj}$ . Для каждого  $\Delta_{kj}$  перенумеруем в лексикографическом порядке все  $C_{2M}^M$  возможных комбинаций отбора. Получим кортеж  $\nu_{k+1} = (\nu_{k+1,j})$  из  $2 \uparrow (m_{k+1} - k - 1)$  номеров. Последовательность  $h = (\nu_1, \nu_2, \dots)$  вектор-номеров  $\nu_k$  взаимно однозначно задает  $\cap D_k$  и тем самым соответствующую меру  $P = P_h\{\cdot\}$ .

Последовательность  $h$  не обладает марковским свойством: чтобы выяснить по номеру  $\nu_k$ , принадлежит ли  $\Delta_k$  конструкции меры  $P_h$ , надо рассмотреть также все предшествующие номера  $\nu_1, \dots, \nu_{k-1}$ . Зато эта последовательность условно независима: нумерации подынтервалов внутри двух разных интервалов одного ранга являются независимыми. Введем в пространстве параметров  $h$  естественную меру, считая на каждом шаге все отборы равновозможными. По теореме Колмогорова о продолжении (см. [3]) эта мера  $\mu$  существует и определяет вероятностную меру на семействе распределений  $\{P_h\}$ , так как нумерация измерима. Обозначим  $\mathfrak{M}$  операцию усреднения по этой мере. Пусть, далее,  $i = (\nu_1, \dots, \nu_k)$  — номер покрытия  $D_{ki}$   $k$ -го ранга. Введем еще усреднение  $\mathfrak{M}_{ki}$  по всем индексам, при которых  $D_{ki}$  является покрытием (т.е. по всем возможным подразбиениям  $D_{ki}$ ), и усреднение  $\mathfrak{M}_{kij}$  по всем занумерованным измельчениям  $j$ -го отрезка  $\Delta_{kij}$ , входящего в  $D_{ki}$ . При  $j \neq l$  усреднения  $\mathfrak{M}_{kij}$  и  $\mathfrak{M}_{kil}$  независимы.

ЛЕММА 2. Пусть  $\Lambda\{\cdot\}$  — лебегова длина на  $E$ ,  $N = N_k$ . Тогда

$$\mathfrak{M}_{kij} P_h\{\cdot | \Delta_{kij}\} = \Lambda\{\cdot | \Delta_{kij}\}, \quad (27)$$

$$\mathfrak{M}_{ki} P_h^N\{\cdot\} - \Lambda^N\{\cdot | D_{ki}\} < (8k(k+1))^{-1}. \quad (28)$$

Доказательство. Каждая из  $2M$  частей  $\Delta_{k+1}$  интервала  $\Delta_{kij}$  участвует ровно в половине отборов, и ей приписывается мера  $M^{-1}P\{\Delta_{kij}\}$ . В другой половине случаев  $\Delta_{k+1}$  приписывается мера нуль. Поэтому

$$\mathfrak{M}_{kij}P_h\{\Delta_{k+1}|\Delta_k\} = \Lambda\{\Delta_{k+1}|\Delta_k\},$$

откуда вытекает (27).

Введем номер  $r$  интервала  $\Delta_{kijr}$  разбиения отрезка  $\Delta_{kij}$  на равные части и обозначим  $\mathbf{j} = (j(1), \dots, j(N))$ ,  $\mathbf{r} = (r(1), \dots, r(N))$ . Тогда

$$\Lambda^N\{ \cdot | D_{ki} \} = \sum_j \sum_{\mathbf{r}} \left[ \frac{1}{2M} P\{\Delta_k\} \right]^N \prod_{s=1}^N \Lambda\{ \cdot | \Delta_{j(s)r(s)} \}, \quad (29)$$

где  $P\{D_{ki}\} = 1$ , поскольку  $\Lambda\{\Delta_{kij}|D_{ki}\} = P\{\Delta_k\}$ . Сходным образом

$$\mathfrak{M}_{ki}P^N\{ \cdot \} = \sum_j \sum_{\mathbf{r}} \mathfrak{M}_{ki} \prod_{s=1}^N [P\{ \cdot | \Delta_{j(s)r(s)} \} P\{\Delta_{j(s)r(s)}\}].$$

Предположим, что все номера  $r(1), \dots, r(N)$  попарно различны (условие А). Тогда все пары  $(j(1), r(1)), \dots, (j(N), r(N))$  заведомо попарно различны, и по условной независимости усреднений

$$\mathfrak{M}_{ki} \prod_s [P\{ \cdot | \Delta_{j(s)r(s)} \} P\{\Delta_{j(s)r(s)}\}] = \prod_s [\Lambda\{ \cdot | \Delta_{j(s)r(s)} \}] \mathfrak{M}_{ki} \prod_s P\{\Delta_{j(s)r(s)}\}.$$

Разобьем  $j(1), \dots, j(N)$  на группы совпадающих индексов. Усреднение  $\mathfrak{M}_{ki}$  действует на каждую группу независимым от других образом. Пусть, например,  $j(1) = \dots = j(q) = j$ . Тогда

$$\mathfrak{M}_{ki} \left[ \prod_{s=1}^q P\{\Delta_{jr(s)}\} \right] = \mathfrak{M}_{kij} \left[ \prod_{s=1}^q P\{\Delta_{jr(s)}\} \right] = \left[ \frac{1}{M} P\{\Delta_k\} \right]^q \frac{C_{2M-q}^{M-q}}{C_{2M}^M}.$$

Значит, при индексах, удовлетворяющих условию А,

$$1 \geq \left[ \frac{1}{2M} P\{\Delta_k\} \right]^{-N} \mathfrak{M}_{ki} \left[ \prod_{s=1}^N P\{\Delta_{j(s)r(s)}\} \right] \geq 1 - \frac{N(N-1)}{2(2M-N+1)}.$$

Наконец, из (29) получаем

$$\Lambda^N\{A|D_{ki}\} = 1(1 - (2M)^{-1}) \dots (1 - (2M - N + 1)^{-1}).$$

Из полученных соотношений вытекает

$$\begin{aligned} |\mathfrak{M}_{ki}P^N\{ \cdot \} - \Lambda^N\{ \cdot | D_{ki} \}| &\leq 2\Lambda^N\{\bar{A}|D_{ki}\} + 2\Lambda^N\{A|D_{ki}\} - 2\mathfrak{M}_{ki}P^N\{A\} \leq \\ &\leq N(N+1)(2M-N+1)^{-1}. \end{aligned}$$

Так как  $2M = 2 \uparrow n_{k+1}$ , где  $n_{k+1}$  определено в (26), то правая часть последнего неравенства меньше  $(8k)^{-1}(k+1)^{-1}$ .

**11.** Вариация разности  $\mathbf{V}[Q : P] = |Q - P|$  неотрицательна и ограничена на  $\text{Cap}(E, \mathcal{L})$  числом 2. Поэтому сильная корректность по вероятности совпадает с  $\mathbf{V}$ -ультрасильной корректностью. Далее, функционал  $\mathbf{V}[Q : P]$  выпукл по  $Q$ . Отсюда по обобщенной теореме Рао–Блекуэлла–Колмогорова (см. [19–21]) усреднение рандомизированной оценки

$$P_N^{\mathbf{x}}\{\cdot\} = \int_{\text{Cap}(E, \mathcal{L})} [Q\{\cdot\}] \Pi(x_1, \dots, x_N; d[Q\{\cdot\}]) \quad (30)$$

только уменьшает  $\mathbf{V}$ -риск.

**ТЕОРЕМА 4.** *Задача статистической оценки неизвестного закона  $P \in \text{Cap}(E, \mathcal{L})$  по независимым наблюдениям без дополнительной априорной информации не является сильно корректной.*

**Доказательство.** Предположим противное. Пусть  $\{\Pi_N\}$  —  $\mathbf{V}$ -ультрасильно состоятельная последовательность детерминированных решающих правил. Построим по ней некоторое специальное семейство сингулярных распределений рассмотренного в предыдущем пункте типа. Пусть  $D_1 = E$ . Число  $N = N_k$  рекуррентно определяется так, чтобы

$$\mathfrak{R}_{\Pi(N)}(\Lambda\{D_{ki}\}) < \frac{1}{4}k(k+1) \quad \forall i = (\nu_1, \dots, \nu_k). \quad (31)$$

Это соотношение вместе с (26) полностью определяет семейство. Из (31) и (28) получаем неравенство

$$\int_0^1 \dots \int_0^1 (\mathfrak{M}_{ki} P^N)\{d\mathbf{x}\} |P_N^{\mathbf{x}}\{\cdot\} - \Lambda\{D_{ki}(P)\}| \leq \frac{1}{2k(k+1)}.$$

Обозначим  $\mathcal{H}_k$  подсемейство сингулярных распределений

$$\left\{ P : \int \dots \int |P_N^{\mathbf{x}}\{\cdot\} - \Lambda\{D_k(P)\}| P^N\{d\mathbf{x}\} < 1 \right\},$$

где  $D_k(P)$  — покрытие носителя  $P$  ранга  $k$ . Переставив<sup>3)</sup> интегрирование по  $d\mathbf{x}$  и усреднение  $\mathfrak{M}$  по мере  $\mu$ , находим

$$\mu\{\mathcal{H}_k\} > 1 - \frac{1}{2k(k+1)}, \quad \mu\{\cap \mathcal{H}_k\} > \frac{1}{2}.$$

Для любого принадлежащего  $\cap \mathcal{H}_k$  распределения  $P_h$  из нашего семейства имеем по определению  $\mathcal{H}_k$  при всех  $k$   $N = N_k$ ,

$$\mathfrak{R}_{\Pi(N)} = \int \dots \int |P_N^{\mathbf{x}}\{\cdot\} - P\{\cdot\}| P^N\{d\mathbf{x}\} \geq 1,$$

так как  $|\Lambda\{D_k(P)\} - P\{\cdot\}| = 2$  ввиду сингулярности  $P$ .

<sup>3)</sup>Перестановочность усреднения  $\mathfrak{M}_{ki}$  и суммирования в интегральной сумме Лебега для измеримой функции  $f(x) = |P_N^{\mathbf{x}}\{\cdot\} - \Lambda\{D_{ki}\}|$  на  $E^N$  очевидна.

**12.** Решающее правило, построенное в доказательстве теоремы 3, требует умения проверять точное совпадение двух наблюдений. Такое требование нельзя считать вполне реалистичным. В математической физике принято (см. [22, 23]) считать задачу корректно поставленной, если ее решение не только существует и единственно, но и устойчиво, т.е. непрерывно зависит от начальных данных (ср. также [24] в связи с постановкой обратных задач). В математической статистике сходным ограничением на класс решающих правил будет требование непрерывной зависимости выводов от результатов наблюдений.

**ОПРЕДЕЛЕНИЕ 4.** Решающее правило  $\Pi_N$  построения оценки по  $N$  наблюдениям

$$\Pi_N : (E^N, \mathcal{L}^N) \rightarrow \text{Сар}(\text{Сар}(E, \mathcal{L}), \mathcal{K}(\mathcal{L}))$$

назовем *непрерывным*, если оно непрерывно отображает куб  $E^N$  в совокупность  $\text{Сар}(\text{Сар}(E, \mathcal{L}), \mathcal{K}(\mathcal{L}))$ , метризованную расстоянием по вариации.

**ЗАМЕЧАНИЕ.** Непрерывное решающее правило  $\Pi_N$  задает, в частности, по (30) непрерывное детерминированное решающее правило  $(x_1, \dots, x_N) \rightarrow P_N^x\{\cdot\}$ , непрерывно отображающее куб  $E^N$  в совокупность  $\text{Сар}(E, \mathcal{L})$ , метризованную расстоянием  $\rho(Q, P) = |Q - P|$ .

**ТЕОРЕМА 5.** *Если ограничиться только непрерывными решающими правилами, то задача статистической точечной оценки для совокупности  $\mathcal{P}_0$  теоремы 3 не будет сильно корректной.*

**Доказательство.** Пусть точка отрезка  $E$ ,  $y \in E$ , есть атом вероятностной меры  $P_N^z\{\cdot\}$ , отвечающей по детерминированному решающему правилу  $\Pi_M : \mathbf{x} \rightarrow P_N^x\{\cdot\}$  кортежу наблюдений  $\mathbf{z} = (z_1, \dots, z_N)$ . Тогда в некоторой окрестности точки  $\mathbf{z}$  также будет  $P_N^x\{y\} > 0$  и, следовательно, будет  $Q_N\{y\} > 0$ , где

$$Q_N\{\cdot\} = \int_0^1 \dots \int_0^1 P_N^x\{\cdot\} dx_1 \dots dx_N.$$

Но мера  $Q_N$  не может иметь более счетного числа атомов. Поэтому найдется  $w \in E$ , не являющееся атомом оценки  $P_N^z$ , ни при одном  $N$  и ни при какой выборке  $\mathbf{z}$ . Следовательно, при  $P^\omega : P^\omega\{\omega\} = 1, |P_N^z - P^\omega| = 2$

$$\mathfrak{R}_{\Pi(N)}(P^\omega) = 2 \quad \forall N.$$

Заметим, что правила построения гистограмм являются лишь кусочно непрерывными. Однако аналог утверждения 3) нашей теоремы 1 сохраняет силу и в классе непрерывных решающих правил. Это вытекает из результатов Нада-рая [15] для оценок типа “ядра”.

**13.** Обратимся теперь к рассмотрению  $\Phi$ -ультрасильной корректности.

**ЛЕММА 3.** *Пусть  $\psi(x) = \varphi(x) - ax + a$ , где  $y = ax - a$  — прикасающаяся прямая к  $\varphi(x)$  при  $x = 1$ . Тогда*

$$\{Q : \Phi_\varphi[Q : P] < \varepsilon\psi(1 - \varepsilon)\} \subseteq \{Q : |Q - P| < 4\varepsilon\}. \quad (32)$$

**СЛЕДСТВИЕ.** *Если задача статистической точечной оценки для априорного семейства  $\mathcal{P}$  является (равномерно)  $\Phi$ -ультрасильно корректной, то она будет также (равномерно) сильно корректной.*

Доказательство. Легко видеть, что  $\Phi_\varphi = \Phi_\psi$ , а выпуклая функция  $\psi(x) > 0$  при  $x \neq 1$  по условию. Очевидно,  $\psi(1 - \varepsilon) \rightarrow 0$  при  $\varepsilon \rightarrow 0$ . Обозначим  $F(z) = P\{x : dQ/dP < z\}$ . Имеем цепочку неравенств (ср., [5, 6])

$$\begin{aligned} \frac{1}{2}|Q - P| &= \int_0^1 (1 - z) dF = \int_0^{1-\varepsilon} (1 - z) dF + \int_{1-\varepsilon}^1 (1 - z) dF \leq \\ &\leq F(1 - \varepsilon) + \varepsilon \leq P\left\{x : \psi\left(\frac{dQ}{dP}\right) > \psi(1 - \varepsilon)\right\} + \varepsilon \leq \\ &\leq [\psi(1 - \varepsilon)]^{-1} \Phi_\psi[Q : P] + \varepsilon \leq 2\varepsilon. \end{aligned}$$

Перейдем к доказательству следствия. Обозначим  $\Pi_N = \Pi(N)$ , тогда

$$(P^{(N)} \circ \Pi_N)\{\Phi[P_N^* : P] \geq \delta\} \leq \delta^{-1} \mathfrak{R}_{\Pi(N)}(P). \quad (33)$$

Поэтому для любой  $\Phi$ -ультрасостоятельной последовательности  $\Pi_N$  имеем в силу (33) и (32) при  $N \rightarrow \infty$  и фиксированном  $\varepsilon > 0$

$$(P_{(N)} \circ \Pi_N)\{|P_N^* - P| > 4\varepsilon\} < [\varepsilon \psi(1 - \varepsilon)]^{-1} \mathfrak{R}_{\Pi(N)}(P) \rightarrow 0.$$

**ТЕОРЕМА 6.** Для априорной совокупности  $\mathcal{P} = \text{Capd}(E, \mathcal{L}, \mathcal{L}_0)$  задача статистической точечной оценки по независимым наблюдениям не является  $\mathbf{I}$ -ультрасильно корректной, где  $\mathbf{I} = \mathbf{I}[Q : P]$  — относительная энтропия.

Доказательство. Предположим, что последовательность решающих правил  $\Pi_N$  состоятельна для меры  $\Lambda\{\cdot\}$  — лебеговой длины на  $(E, \mathcal{L})$ . Так как  $\mathbf{I}[Q : P] = +\infty$ , если  $P$  не доминирует  $Q$ , то, начиная с некоторого  $N$ , необходимо

$$(\Lambda^N \circ \Pi_N)\{\Lambda \gg P_n^*\} = 1,$$

откуда при всех  $P \ll \Lambda$

$$(P^N \circ \Pi_N)\{\Lambda \gg P_n^*\} = 1.$$

Пусть  $P \gg Q$ ,  $G = \text{supp}_P Q = \{x : 0 < dQ/dP\}$ ,  $P'\{\cdot\} = P\{\cdot|G\}$ . Тогда (см. [2])

$$\mathbf{I}[Q : P] = \mathbf{I}[Q : P'] - \ln P\{G\} \geq -\ln P\{G\}. \quad (34)$$

Из (33) и (34) вытекает, что, начиная с некоторого  $N = N_0$ ,

$$(\Lambda^N \circ \Pi_N)\left\{P_N^* : \Lambda\{\text{supp}_\Lambda P_N^*\} \geq \frac{3}{4}\right\} \geq \frac{1}{2}.$$

Отсюда при каждом  $N \geq N_0$  найдется куб

$$\Delta = \Delta^{(N)} = \Delta_1^{(N)} \times \dots \times \Delta_N^{(N)}$$

с двоично-рациональными вершинами и ребром  $r_N < N^{-1}2^{-N-2}$ , для которого

$$\frac{1}{r_N^N} \int \dots \int_{\Delta} d\mathbf{x} \Pi(x_1, \dots, x_N; [Q : \Lambda\{\text{supp}_\Lambda Q\} \geq \frac{3}{4}]) \geq \frac{1}{2}. \quad (35)$$

Пусть

$$H_N = \Delta_1^{(N)} \cup \dots \cup \Delta_N^{(N)}, \quad H = \cup H_N, \quad N \geq N_0.$$

Тогда

$$\Lambda\{H\} \leq \frac{1}{2}, \quad P\{\cdot\} = \Lambda\{\cdot|H\} \ll \Lambda\{\cdot\}, \quad \Lambda\{\text{supp}_\Lambda P\} = \Lambda\{H\} \leq \frac{1}{2}.$$

Из (35) вытекает, что

$$(P^N \circ \Pi_N) \left\{ P_N^* : \Lambda\{\text{supp}_\Lambda P_N^*\} \geq \frac{3}{4} \right\} = c_N > 0.$$

Это означает, что с положительной  $P^N$ -вероятностью (не меньшей  $c_N$ ) оценка  $P_N^*$  не доминируется теоретическим законом  $P$ , откуда  $\mathfrak{R}_{\Pi(N)}(P) = +\infty$ .

**14.** Из наших теорем следует четкий вывод: без дополнительной априорной информации задача статистической оценки по наблюдениям неизвестного закона распределения на нетривиальном измеримом пространстве поставлена некорректно.

Наиболее важной корректно поставленной задачей статистической точечной оценки является параметрическая задача, когда априорное семейство  $\mathcal{P}$  является гладким семейством  $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$  от векторного параметра  $\theta$ , изменяющегося в некоторой области  $\Theta \subseteq \mathbb{R}^n$ . Здесь точность оценок ограничивается снизу известным неравенством информации, устанавливающим границу различимости близких законов по выборке. При достаточно широких предположениях равномерная I-ультрасильная корректность параметрических задач и точная асимптотика минимакса риска были установлены нами в [2], что несколько уточняет классические результаты (см. [8]) в этой области.

Естественно выделяется также класс счетно параметрических задач, когда априорное семейство  $\mathcal{P}$  гладко описывается счетным числом параметров, но для задания  $P \in \mathcal{P}$  с фиксированной точностью достаточно лишь конечного числа из них. Например, априорная информация о наличии у  $P$  достаточно гладкой плотности выделяет семейство  $\mathcal{P}$  указанного типа. Наиболее интересны (см. [25]) компактные экспоненциально выпуклые семейства  $\mathcal{P}$ , хорошо аппроксимируемые экспонентными семействами. Равномерная I-ультрасильная корректность задачи статистической точечной оценки для такого семейства при условии убывания его  $n$ -мерных информационных поперечников Колмогорова была доказана в [2]. Заметим, что впервые развитый в [26] систематический подход к выпуклой счетно параметрической задаче статистической точечной оценки на основе концепции поперечников [27] позволяет строить для этой задачи решающие правила, оптимальные по порядку убывания риска.

## СПИСОК ЛИТЕРАТУРЫ

1. Ченцов Н.Н. О статистических оценках неизвестного закона распределения вероятностей // Теория вероятн. и ее примен. 1980. Т. 25, вып. 3. С. 647–648.
2. Ченцов Н.Н. Статистические решающие правила и оптимальные выводы. — М.: Наука, 1972.
3. Колмогоров А.Н. Основные понятия теории вероятностей. — М.: ОНТИ, 1936.
4. Прохоров Ю.В. О случайных мерах на компакте // Докл. АН СССР. 1961. Т. 138, № 1. С. 53–55.
5. Csizsar I. Information-type indices of the divergence of distributions. I, II // Magyar Tud. Akad. Mat. 1967. V. 17, № 2. P. 123–149; № 3, P. 267–291.

6. *Csiszar I.* On topological properties of  $f$ -divergence // *Studia Sci. Math. Hungar.* 1967. V. 2, № 3–4. P. 329–339.
7. *Cencov N.N.* Algebraic foundations of mathematical statistics // *Math. Operat. Statist., Ser. Statist.* 1978. V. 9, № 2. P. 267–276.
8. *Cramèr H.* *Mathematical methods of statistics.* — Princeton Univ. Press, 1946.
9. *Wald A.* Contributions to the theory of statistical estimation and testing hypotheses // *Ann. Math. Statist.* 1939. V. 10, № 4. P. 299–326.
10. *Le Cam L.* On some asymptotic properties of maximum likely hood estimation and related Bayes estimates // *Univ. California Publ. Statist.* 1953. V. 1. P. 267–329.
11. *Levy P.* *Théorie de l'addition des variables aléatoires.* — Paris: Gauthier-Villars, 1937.
12. *Прохоров Ю.В.* Сходимость случайных процессов и предельные теоремы теории вероятностей // *Теория вероятн. и ее примен.* 1956. Т. 1, вып. 2. С. 177–235.
13. *Гливенко В.И.* Sulla determinazione empirica di una legge di probabilita // *Giornale dell'Istituto Italiano degli Attuari.* 1933 V. 4, № 1. P. 1–10.
14. *Колмогоров А.Н.* Sulla determinazione empirica di una legge di distribuzione // *Giornale dell'Istituto Italiano degli Attuari.* 1933. V. 4, № 1. P. 83–91.
15. *Надарая Э.А.* О непараметрической оценке байесовского риска в задаче классификации // *Сообщ. АН ГрузССР.* 1976. Т. 82, № 2. С. 277–280.
16. *Колмогоров А.Н., Фомин С.В.* *Элементы теории функций и функционального анализа.* — М.: Наука, 1976.
17. *Колмогоров А.Н.* Оценки минимального числа  $\varepsilon$ -сетей в различных функциональных пространствах и их применение к вопросу о представимости функций нескольких переменных суперпозициями функций меньшего числа переменных // *Успехи матем. наук.* 1955. Т. 10, вып. 1. С. 192–193.
18. *Колмогоров А.Н., Тихомиров В.М.*  $\varepsilon$ -энтропия и  $\varepsilon$ -емкость множеств в функциональных компактах // *Успехи матем. наук.* 1959. Т. 14, вып. 2. С. 3–86.
19. *Rao C.R.* Information and the accuracy attainable in the estimation of statistical parameters // *Bull. Calcutta Math. Soc.* 1945. V. 37, № 3. P. 81–91.
20. *Blackwell D.* Conditional expectation and unbiased sequential estimation // *Ann. Math. Statist.* 1947. V. 18, № 1. P. 105–110.
21. *Колмогоров А.Н.* Несмещенные оценки // *Изв. АН СССР. Сер. матем.* 1950. Т. 14, № 4. С. 303–326.
22. *Hadamard J.* *Le problème de Cauchy.* — Paris: Hermann, 1932.
23. *Courant R., Hilbert D.* *Methoden der mathematische Physik. II.* — Berlin: Springer, 1937.
24. *Тихонов А.Н.* Об устойчивости обратных задач // *Докл. АН СССР.* 1943. Т. 39, № 5. С. 195–198.
25. *Стратонович Р.Л.* Быстрота сходимости алгоритмов оценки плотности распределения вероятностей // *Изв. АН СССР. Сер. техн. киберн.* 1969. Т. 6. С. 3–15.
26. *Ченцов Н.Н.* Оценка неизвестной плотности распределения по наблюдениям // *Докл. АН СССР.* 1962. Т. 147, № 1. С. 45–48.
27. *Kolmogoroff A.* Über die besste Annäherung von Funktionen einen gegebenen Funktionenklasse // *Ann. Math.* 1936. V. 37, № 1. P. 107–110.
28. *Abou-Jaoude S.* Conditions nécessaire et suffisantes de convergence  $L_1$  en probabilité l'histogramme pour une densité // *Ann.Inst. Henri-Poincaré. Sect. B.* 1976. V. 12, № 3. P. 213–231.

## ПОЧЕМУ $L_1$ -ПОДХОД И ЧТО ЗА ГОРИЗОНТОМ? \*)

Монография Л. Девроя и Л. Дьёрфи посвящена задаче статистической оценки плотности распределения по наблюдениям. В книге сделан обзор современного состояния исследований в этой области, доведенный в ряде мест до 1983 г., подробно рассмотрены наиболее интересные и важные результаты, а также вопросы, остающиеся нерешенными. Название книги точно отражает ее содержание, так что неискушенному читателю может показаться, что книга двух сравнительно молодых специалистов посвящена узкому, даже экзотическому специальному вопросу. Между тем  $L_1$ -подход является наиболее простым и естественным в рассматриваемой проблеме, и можно дать тому строгое математическое доказательство. Для этого надо лишь обратиться к общей теории статистических выводов, заложенной А. Вальдом [1]. В предлагаемом ниже кратком очерке некоторых основных понятий математической статистики выясняется место рассмотренной в книге проблематики и тем самым дополняются и уточняются суждения авторов на этот счет (в частности, по поводу  $L_1$ -подхода и его сравнения с  $L_2$ -подходом).

В теории вероятностей случайное явление задается своим вероятностным пространством  $(\Omega, \mathcal{A}, P)$ . В этой идеализированной схеме  $\Omega$  — пространство всех мыслимых исходов  $\omega \in \Omega$  эксперимента,  $\mathcal{A}$  есть  $\sigma$ -алгебра подмножеств  $A \subseteq \Omega$ , называемых событиями,  $P$  — вероятностная мера на  $\mathcal{A}$ , называемая распределением вероятностей исходов. Измеримое пространство  $(\Omega, \mathcal{A})$  задает качественное описание случайного явления, а мера  $P$  — количественное. Пусть  $\omega^1, \dots, \omega^N$  — последовательность независимых наблюдений данного явления, т.е.  $P^N\{d\omega\} = P\{d\omega^1\} \dots \{d\omega^N\}$ . Тогда по закону больших чисел для любого события  $A \in \mathcal{A}$  частота  $\nu(A)/N \approx P\{A\}$ , где  $\nu_N(A) = \text{Card}\{\omega^k: \omega^k \in A, k = 1, \dots, N\}$ , причем точность и надежность (по  $P^N$ -вероятности) этого предсказания растут с увеличением  $N$ . Задача предсказания частот является одной из основных в теории вероятностей и во многом определяет ее практическую ценность. В математической статистике мы сталкиваемся с обратной задачей: нам дана последовательность наблюдений  $\omega^1, \dots, \omega^N$  некоторого случайного явления с известным качественным описанием  $(\Omega, \mathcal{A})$  и требуется оценить “наблюдаемое” распределение  $P$  или какую-то его характеристику либо выяснить наличие у этого неизвестного наблюдателю закона  $P$  тех или иных свойств. В первом случае говорят о задаче *статистической точечной оценки* (с.т.о.), и именно одна из ее постановок и составляет предмет исследования Девроя и Дьёрфи, а также нашего рассказа. Здесь только следует добавить, что при решении иных статистических задач процедура точечного оценивания закона  $P$  используется как удобный промежуточный этап, когда о свойствах  $P$  судят по построенной его оценке  $P^*$ .

Обратные задачи математической физики, как правило, некорректны. Не составляет в этом отношении исключения и обратная задача теории вероятностей.

\*) Деврой Л., Дьёрфи Л. Непараметрическое оценивание плотности.  $L$ -подход. — М.: Мир, 1988. — С. 348–361.

Для ее корректной сильной разрешимости нужна дополнительная априорная информация. Например, может быть заранее известно, что наблюдаемое  $P$  принадлежит семейству  $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$ . Когда “координатизирующее” множество  $\Theta$  конечно, задачу статистической точечной оценки называют *задачей проверки* (нескольких простых) *гипотез*. Когда законы  $P_\theta$  гладко зависят от конечномерного векторного параметра  $\theta$ , говорят о *параметрической задаче оценивания*. Основное внимание в этой книге уделяется случаю, когда такое априорное семейство  $\mathcal{P}$  законов может быть гладко запараметризовано только счетным числом вещественных координат. Так будет, например, если известно лишь, что наблюдаемая случайная величина ограничена, а ее распределение имеет плотность  $p(x)$ , принадлежащую функциональному пространству  $C^{(2)}$ . Наконец, семейство  $\mathcal{P}$  может быть столь обширно, что оно не допускает даже счетномерной гладкой параметризации. Так будет, например, когда известно лишь, что координаты наблюдаемой двумерной случайной величины независимы (друг от друга), и ничего более. По традиции обе последние постановки объединяют под названием *задачи непараметрического оценивания*, хотя они резко отличаются и по подходам, и по методам решения. В частности, последняя задача корректна лишь в слабом смысле. Приведенный перечень не дает, разумеется, исчерпывающей классификации всех возможных постановок задачи с.т.о. Мы указали лишь наиболее важные из них, расположив их в порядке убывания априорной информации и вытекающего отсюда убывания точности решения. Так, при проверке простых гипотез вероятность ошибки убывает экспоненциально, в конечно-параметрической задаче с.т.о. погрешность в значении параметра и закона имеет порядок  $N^{-1/2}$ , а в счетно-параметрической порядок убывания не достигает и этой величины. Как мы уже отмечали, именно последняя задача и связанные с нею аспекты конечно-параметрической задачи являются предметом нашего рассмотрения.

Согласно догме Вальда [1], в каждой статистической задаче кроме вводных данных должно быть указано измеримое пространство  $(\Delta, \mathcal{B})$  выводов  $\delta \in \Delta$ , к одному из которых мы должны прийти в результате статистического анализа данных эксперимента. Тогда каждое детерминированное решающее правило (т.е. избранная статистиком процедура обработки наблюдений) задается своей функцией  $\delta = f(\omega^1, \dots, \omega^N)$ , а стохастическое (или рандомизированное) — функцией  $\delta = f(\omega^1, \dots, \omega^N; \eta)$ , где  $\eta$  — случайный параметр, который приходится дополнительно разыгрывать. Теоретически рандомизированные (и даже детерминированные) решающие правила удобнее описывать неявно, переходным распределением вероятностей  $\mathbb{I}(\omega^1, \dots, \omega^N; d\delta)$ , не фиксируя способа конструирования дополнительного случайного эксперимента, в котором будет “разыгран” вывод.

Точность выбранного решения  $\delta$  Вальд предлагает определять функцией потерь  $\mathcal{L}(\theta, \delta)$  — убытком, который терпит статистик, принявший решение  $\delta$ , когда на самом деле он наблюдает явление  $P_\theta$ . Тогда качество решающего правила  $\mathbb{I}$  можно охарактеризовать математическим ожиданием потерь — риском

$$\mathcal{R}_{\mathbb{I}(N)} = E\mathcal{L}(\theta, \delta) = \int_{\Delta} (P_\theta^N \mathbb{I})\{d\delta\} \mathcal{L}(\theta, \delta), \quad (1)$$

где вероятностная мера  $P_{\theta, \Pi}^{(N)} = P_{\theta}^N \Pi$  задается, как

$$P_{\theta, \Pi}^{(N)}\{\cdot\} = \int_{\Omega^N} \Pi(\omega^1, \dots, \omega^N; \cdot) P_{\theta}\{d\omega^1\} \dots P_{\theta}\{d\omega^N\}. \quad (2)$$

Заметим в скобках, что вместо риска можно рассматривать какой-либо квантиль случайных потерь. Так или иначе, зная семейство  $\mathcal{P} = \{P_{\theta}\}$ , мы можем заранее вычислять распределения  $P_{\theta, \Pi}^{(N)}$  и подбирать по ним наиболее “выгодное” решающее правило  $\Pi$ .

В обратной задаче теории вероятностей пространством  $\Delta$  всевозможных ответов  $P^*$  является в принципе вся полная совокупность  $\text{Car}(\Omega, \mathcal{A})$  всех распределений вероятностей  $P$  на  $(\Omega, \mathcal{A})$ . Для лебеговых измеримых пространств  $(\Omega, \mathcal{A})$ , преодолев, следуя Ю.В. Прохорову, некоторые технические трудности (см. [2, § 29]), можно снабдить эту совокупность  $\sigma$ -алгеброй событий и построить измеримое пространство случайных вероятностных мер  $\text{Car}(\text{Car}(\Omega, \mathcal{A}), \mathcal{K}(\mathcal{A}))$ . При этом параметрами закона  $P$  служат значения  $P\{A_i\}$  на некоторой системе образующих  $A_i$  алгебры  $\mathcal{A}$  и зависимость  $P$  от параметров  $\mathcal{K}(\mathcal{A})$  измерима, но не гладка.

В качестве величины погрешности оценки  $P^*$  искомой вероятностной меры  $P$  естественно в первую очередь взять вариацию их разности

$$|P^* - P| = \sup_{A \in \mathcal{A}} [P^*\{A\} - P\{A\}] - \inf_{A \in \mathcal{A}} [P^*\{A\} - P\{A\}], \quad (3)$$

т.е. задать ее через сильную норму в линейном пространстве  $L_1(\Omega, \mathcal{A})$  всех (знакопеременных, вообще говоря) счетно-аддитивных мер на  $(\Omega, \mathcal{A})$ ;  $L_1(\Omega, \mathcal{A}) = \text{Lin Car}(\Omega, \mathcal{A})$ .

Следующий точный результат принадлежит автору этих строк (см. [3, теорему 4]).

**ТЕОРЕМА 1.** *Когда погрешность оценки с.т.о. измеряется сильной нормой, задача с.т.о. для совокупности  $\text{Car}(E, \mathcal{A}^*)$ , где  $E$  — единичный отрезок, а  $\mathcal{A}^*$  — алгебра всех его абсолютно измеримых лебеговых подмножеств, без дополнительной априорной информации некорректна.*

Оказывается, что для любой последовательности решающих правил

$$\Pi(N): \text{Car}(E^N, \mathcal{A}^{*N}) \rightarrow \text{Car}(\text{Car}(E, \mathcal{A}^*), \mathcal{K}(\mathcal{A}^*))$$

существует такой вероятностный закон  $P \in \text{Car}(E, \mathcal{A}^*)$ , что

$$\lim_{N \rightarrow \infty} \mathcal{R}_{\Pi(N)}(P) \geq 1. \quad (4)$$

Наряду с сильной на совокупности  $\text{Car}(E, \mathcal{A}^*)$  и ей подобными существуют и слабые метрики, выражающиеся через функции распределения. Простейшая из них задается  $C$ -расстоянием

$$\rho(P, Q) = \sup_x |F(x) - G(x)|,$$

где  $F(x) = P\{[0, x)\}$ ,  $G(x) = Q\{[0, x)\}$ . По известной теореме Гливенко [4] при  $N \rightarrow \infty$

$$\sup_x |F_N^*(x) - F(x)| \rightarrow 0$$

почти наверно, так что эмпирическая функция распределения  $F_N^*$  является состоятельной оценкой для теоретической  $F$ . Таким образом, некорректность обратной задачи теории вероятностей не слишком сильна — примерно такая же, как и у операции численного дифференцирования в вещественном анализе.

Мы не будем здесь выяснять, насколько приближения в слабой метрике менее информативны, чем приближения в сильной метрике, отсылая за этим читателя к [5]. Ограничимся примером. Как известно, всякая мера на прямой может быть разложена в линейную комбинацию мер трех типов — дискретной, непрерывной и сингулярной. Мера первого типа сосредоточена на не более чем счетном множестве точек, имеющих каждая положительную меру, мера второго типа имеет плотность (по лебеговой длине), а мера третьего типа, не имея точек положительной меры, сосредоточена на каком-то подмножестве лебеговой длины нуль. Эти три типа мер резко различаются по виду функции распределения. Разделены они между собой и в сильной метрике — вариация разности двух мер разных типов равна 2. А в слабых метриках непрерывный и сингулярный типы не разделены. И несколько препарируя доказательство теоремы 1, можно показать, что не существует состоятельных решающих правил, позволяющих по возрастающей выборке определить, непрерывен или сингулярен “наблюдаемый” закон  $P$ . Любопытно отметить, что отличить дискретный тип от остальных или выделить дискретную компоненту можно, только обладая способностью устанавливать абсолютно точное совпадение двух результатов эксперимента, и невозможно в противном случае.

Из приведенных соображений нетрудно вывести, что максимальным семейством, для которого задача с.т.о. может иметь смысл в сильной метрике, является подсовокупность всех доминированных мер  $\text{Card}(\Omega, \mathcal{A}, Z)$  на каком-либо измеримом пространстве  $(\Omega, \mathcal{A})$  с фиксированным идеалом  $Z$  нуль-множеств, в частности, подсовокупность всех распределений вероятностей на вещественной прямой или единичном отрезке, имеющих плотность (по лебеговой длине). Для таких подсовокупностей расстояние по вариации превращается в обычное  $L_1$ -расстояние между плотностями распределений:

$$|Q - P| = \int_{\Omega} |q(\omega) - p(\omega)| \mu\{d\omega\}, \quad (5)$$

где  $\mu$  — доминирующая мера, например,  $\mu\{dx\} = dx$  для единичного отрезка, а оценка вероятностного закона сводится к оценке его плотности. Корректность такой постановки задачи с.т.о. была установлена в 1976 г. независимо Абу-Жаудом [6] (его результаты приведены Девроем и Дьёрфи) и Э.А. Надарая [7]. Первый в качестве оценки плотности использовал гистограммы, а второй — оценки типа ядра. Из инвариантности подсовокупности всех распределений вероятностей непрерывного типа относительно взаимно измеримых сохраняющих длину отображений отрезка на себя легко вытекает (см. [3]), что универсальное состоятельное решающее правило не может быть на этой подсовокупности равномерно состоятельным. Однако для более узких априорных семейств  $\mathcal{P}$  законов

те же решающие процедуры могут приводить к равномерно на всем семействе  $\mathcal{P}$  убывающему риску.

Дело теоретика — искать в первую очередь наиболее общие закономерности, наиболее общие подходы. Попытаемся сейчас выяснить, чем же  $L_2$ -подход в этом отношении хуже  $L_1$ -подхода.

Каждое статистическое решающее правило  $\Pi$  задает аффинное отображение совокупности  $\text{Car}(\Omega^N, \mathcal{A}^N)$  в  $\text{Car}(\Delta, \mathcal{B})$ . Это связано с тем и только тем, что правило  $\Pi$  задается переходным распределением вероятностей. Рассмотрим все возможные объекты  $\text{Car}(\Omega, \mathcal{A})$  и все возможные переходные распределения вероятностей  $\Pi$  из какого-либо одного измеримого пространства  $(\Omega, \mathcal{A})$  в какое-то другое  $(\Omega', \mathcal{A}')$ . Они образуют алгебраическую категорию  $\text{CAR}$ , объектами которой служат совокупности  $\text{Car}(\Omega, \mathcal{A})$ , а морфизмы (здесь это марковские отображения) задаются переходными распределениями вероятностей по правилу

$$Q\{\cdot\} = (P\Pi)\{\cdot\} = \int_{\omega} P\{d\omega\}\Pi(\omega; \cdot).$$

При этом выполнены все аксиомы категории.

1. Тожественное отображение  $\Pi$  каждого объекта на себя принадлежит категории  $\Pi(\omega; B) = \chi_B(\omega)$ , где  $\chi_B(\cdot)$  — индикатор множества  $B$ .

2. Композиция  $\Pi \circ \mathcal{K}$  двух марковских отображений снова есть марковское отображение  $\Pi$ :

$$\Pi(\omega; \cdot) = \int_{\Omega'} \Pi(\omega; d\omega') \mathcal{K}(\omega'; \cdot).$$

3. Композиция является ассоциативной операцией,  $\Pi(\mathcal{K}\mathcal{C}) = (\Pi\mathcal{K})\mathcal{C}$ , как композиция отображений.

Этот факт был впервые отмечен Н.Н. Ченцовым [8] и независимо Н. Морсом и Р. Сакстедером [9]. Кроме того, в этой категории  $\text{CAR}$  имеются дополнительная операция умножения объектов и связанные с ней операции умножения и усреднения морфизмов [10]. Отсюда нетрудно установить (см. [10, 11]), что два семейства распределений вероятностей  $\{P_{\theta}, \theta \in \Theta\}$  и  $\{Q_{\theta}, \theta \in \Theta\}$ , координатизированные одним и тем же множеством  $\Theta$ , обладают одинаковыми статистическими свойствами тогда и только тогда, когда существуют такие два марковских морфизма  $\Pi$  и  $\mathcal{K}$ , что

$$P_{\theta}\Pi = Q_{\theta}, \quad Q_{\theta}\mathcal{K} = P_{\theta} \quad \forall \theta \in \Theta.$$

Во всякой теории общий закон должен допускать эквивариантную формулировку. Другими словами, его утверждение не должно изменяться при переходе от одной ситуации к другой, эквивалентной первой (эквивалентной в рамках этой теории), иначе он не будет общим законом. В классической геометрии такие переходы образуют группу, в теории статистического вывода — категорию, порождающую своеобразную геометрию. Семейства распределений играют в ней роль “фигур”, а марковские морфизмы — “движений”. И многие основные понятия математической статистики получают истолкования как инварианты, коварианты или более сложные эквиварианты этой геометрии. Так как марковские

морфизмы, вообще говоря, необратимы, кроме обычного для групповой геометрии понятия инварианта — величины, принимающей одинаковые значения для конгруэнтных фигур (т.е. статистически эквивалентных семейств) — возникает понятие монотонного инварианта. Нам оно понадобится только для случая пары вероятностных законов; определение для случая более обширных семейств совершенно аналогично.

ОПРЕДЕЛЕНИЕ. Функцию  $f(P, Q)$  от двух аргументов, определенную на квадратах всех объектов  $\text{Car}(\Omega, \mathcal{A})$ , назовем *монотонным инвариантом*, если

$$f(P\mathbb{I}, Q\mathbb{I}) \leq f(P, Q) \quad (6)$$

при всех допустимых наборах  $P, Q, \mathbb{I}$  аргументов.

Примерами монотонных инвариантов от пары распределений могут служить вариация разности  $|P - Q|$  (см. (3)), относительная энтропия

$$\mathfrak{E}(P, Q) = \int_{\Omega} \left[ \frac{dQ}{dP}(\omega) \ln \frac{dQ}{dP}(\omega) \right] P\{d\omega\} = \int_{\Omega} \left[ \ln \frac{dQ}{dP}(\omega) \right] Q\{d\omega\},$$

расстояние Бхаттачария

$$s(P, Q) = 2 \arccos \int_{\Omega} \sqrt{P\{d\omega\} Q\{d\omega\}}, \quad (7)$$

порождаемое квадратичной формой Фишера (фишеровской информацией). Кстати, это расстояние — единственная с точностью до постоянного множителя инвариантная риманова метрика на объектах категории  $\text{CAR}$ . И говоря о гладких конечномерных семействах  $\{P_{\theta}, \theta \in \Theta\}$ , мы почти всегда молчаливо предполагаем, что функция  $P_{\theta}$  аргумента  $\theta$  в области  $\Theta \supseteq R^n$  дифференцируема в смысле этой метрики. К классу монотонных инвариантов относятся, как нетрудно установить, и  $\varphi$ -дивергенции Чисара [12, 13]

$$I_{\varphi}(P, Q) = \int \varphi\left(\frac{dP}{dQ}(\omega)\right) Q\{d\omega\},$$

где  $\varphi$  — выпуклая функция на  $R^+$ . Для взаимно абсолютно непрерывных мер определение энтропии упрощается до

$$\mathfrak{E}(P, Q) = \int_{\Omega} \left[ \ln \frac{q(\omega)}{p(\omega)} \right] q(\omega) \mu\{d\omega\}, \quad (8)$$

а форма Фишера на семействе  $\{P_{\theta}\}$  записывается как

$$ds^2 = \sum_{\alpha, \beta} d\theta_{\alpha} d\theta_{\beta} \int_{\Omega} \frac{\partial \ln p(\omega; \theta)}{\partial \theta_{\alpha}} \cdot \frac{\partial \ln p(\omega; \theta)}{\partial \theta_{\beta}} P_{\theta}\{d\omega\}. \quad (9)$$

Исключительная роль  $L_1$ -нормы в классе инвариантных метрик устанавливается следующей теоремой; см. [14].

ТЕОРЕМА 2. Если метрика  $\rho$  на объектах категории  $\text{CAR}$  монотонна в категории, то

$$\rho(P, Q) \geq \frac{1}{8} \rho(R_{1/2}, R_{1/4}) \cdot |P - Q|, \quad (10)$$

где  $R_\theta$  — распределения вероятностей на  $\Omega_2 = \{\omega_1, \omega_2\}$ ,

$$R_\theta\{\omega_1\} = \theta, \quad R_\theta\{\omega_2\} = 1 - \theta, \quad 0 \leq \theta \leq 1.$$

Таким образом, если статистическая задача некорректна в расстоянии по вариации, то она остается некорректной и при измерении величины погрешности любой другой инвариантной метрикой (обратное неверно). В то же время (см. [2]) расстояние по вариации является единственным (с точностью до множителя) инвариантным расстоянием, определяемым нормой разности. Точнее, если метрика  $\rho$  на объектах категории CAP инвариантна относительно категории и однородна, т.е.

$$P - Q = \lambda(P' - Q') \implies \rho(P, Q) = \lambda \cdot \rho(P', Q'),$$

то

$$\rho(P, Q) = \frac{1}{2} \rho(R_{1/2}, R_{1/4}) \cdot |P - Q|.$$

Сходная ситуация имеет место и для функций потерь самого общего вида (для  $\varphi$ -дивергенций это было доказано Чисаром), лишь только бы они являлись монотонными инвариантами; см. [14].

ТЕОРЕМА 3. Для монотонного инварианта  $\mathcal{L}(P, Q)$  со свойством

$$Q \neq P \implies \mathcal{L}(P, Q) \neq \mathcal{L}(P, P)$$

существуют константа  $c = \mathcal{L}(P, P) \quad \forall P$  и монотонная числовая функция

$$g(z) = \inf_{2|x-y| \geq z} \rho(R_x, R_y), \quad (11)$$

где  $0 \leq z \leq 2$ ,  $g(z) > 0$  при  $z > 0$ ,  $g(0) = 0$ , такие, что

$$\mathcal{L}(P, Q) \geq c + g(|P - Q|).$$

Отсюда вытекает, что  $L_1$ -подход, излагаемый авторами этой книги, является наиболее общим инвариантным подходом в статистической задаче оценки плотности, что мы и обещали строго доказать.

Что же касается  $L_2$ -подхода, который также рассматривают Деврой и Дьёрфи, то он неинвариантен, поскольку квадрат нормы

$$\int_{\Omega} [p(\omega) - q(\omega)]^2 \mu\{d\omega\} \quad (12)$$

уже зависит от выбора доминирующей меры  $\mu$ . И вполне может случиться, что при замене меры конечная норма разности превратится в бесконечную (в то же время характеристики (5) и (8) близости двух вероятностных законов инвариантны относительно этого выбора). Тем не менее для некоторых классов априорных семейств  $\mathcal{P}$  удастся развить квазиинвариантный подход, использующий целый “букет” эквивалентных  $L_2(P)$ -норм. Этот подход будет указан несколько далее.

Несмотря на неинвариантность,  $L_2$ -подход удобен тем, что позволяет строить простые алгоритмы восстановления неизвестной плотности по независимым наблюдениям и легко характеризовать их точность в выбранной  $L_2$ -норме. С ним, например, естественно связаны так называемые *проекционные оценки плотности* (см. гл. 12), предложенные впервые нами в 1958 г. и исследованные затем в [15, 16]. Как показали эксперименты В.В. Статулявичюса [17], при их использовании обработка данных может идти на порядок быстрее, чем при использовании ядерных оценок Розенבלата–Парзена с фиксированной формой ядра, а результаты обработки требуют на порядок меньшего объема памяти для хранения. Но они, как и всякие ядерные методы со знакопеременными ядрами, обладают существенным недостатком. Отрезок ряда Фурье с подставленными оценками коэффициентов может не оказаться неотрицательной функцией, т.е. построенная проекционная оценка не будет плотностью вероятностной меры. Разумеется, можно заменить ее на участках отрицательности нулем и отнормировать к единичному интегралу. Как показывают Деврой и Дьёрфи (см. с. 277),  $L_1$ -погрешность при этом только уменьшится. Но при такой дополнительной процедуре теряются и простота конструкции, и простота хранения. Возникает естественное желание построить метод, лишенный указанного недостатка, достаточно точный и к тому же эквивариантный. Заметим, что требование точности является существенным. Как известно, полигон частот, построенный по гистограмме, сходится по вероятности к искомой гладкой плотности класса  $C^{(2)}$  со скоростью  $N^{-2/5}$ . Таков же порядок погрешности ядерных оценок с неотрицательными ядрами ввиду их смещенности.

Отмеченных выше недостатков лишен экспонентный метод оценивания плотности, предложенный Р.Л. Стратоновичем [18] и независимо автором этих строк [2]. Идея состоит в том, чтобы приближать линейным агрегатом не саму плотность, а ее логарифм. Поскольку интеграл от функции  $p_0(\omega) \exp[s^j q_j(\omega)]$ , где по совпадающим верхнему и нижнему индексам предполагается суммирование, не обязан быть тождественно равен единице, эту функцию надо отнормировать, что дает формулы

$$p(\omega; \mathbf{s}) = p_0(\omega) \exp[s^j q_j(\omega) - \Psi(\mathbf{s})], \quad (13)$$

$$\Psi(\mathbf{s}) = \ln \int_{\Omega} \exp[s^j q_j(\omega)] p_0(\omega) \mu\{d\omega\}, \quad (14)$$

где  $\psi(\mathbf{s})$  — логарифм нормирующего делителя. Формула (13) имеет смысл, когда интеграл в (14) конечен. Нетрудно показать, что функция  $\psi(\mathbf{s})$  выпуклая, так что область  $\{\mathbf{s}: \psi(\mathbf{s}) < +\infty\}$  выпукла (хотя и может быть пустой).

Таким образом, мы построили экспонентное семейство  $\gamma$  распределений вероятностей с каноническим параметром  $\mathbf{s}$  и областью задания  $\text{Dom } \gamma = \{\mathbf{s}: \psi(\mathbf{s}) < +\infty\}$ . Естественно, что каноническая параметризация определена с точностью до аффинного преобразования параметров и направляющих статистик  $q_j(\omega)$ .

Экспонентные семейства возникают в очень многих вопросах математической статистики (и статистической физики). За их последовательной теорией мы отсылаем к [2, 19, 20]. Введем вектор-функцию

$$\mathbf{t} = \mathbf{T}(\mathbf{s}) = \int_{\Omega} \mathbf{q}(\omega) p(\omega; \mathbf{s}) \mu\{d\omega\}. \quad (15)$$

Хорошо известно, что  $\mathbf{T}(\mathbf{s}) = \text{grad } \psi(\mathbf{s})$  — взаимно однозначная и аналитическая вектор-функция внутри области  $\text{Dom } \gamma$ , являющаяся преобразованием Лежандра, так что

$$\mathbf{s} = \text{grad} [s^j(\mathbf{t})t_j - \Psi(\mathbf{s}(\mathbf{t}))].$$

Несколько более сложны связи  $\mathbf{s}$  и  $\mathbf{t}$  на границе области  $\text{Dom } \gamma$ ; см. [2].

Параметр  $\mathbf{t}$  мы будем называть *натуральным параметром экспонентного семейства*. Для него существует простая несмещенная статистическая оценка

$$\mathbf{t}^* = N^{-1}[\mathbf{q}(\omega^1) + \dots + \mathbf{q}(\omega^N)]. \quad (16)$$

Эта оценка эффективна, т.е. неравенство информации для нее обращается в равенство. Отметим, что справедливо и обратное: если параметр семейства допускает эффективную оценку, то это семейство является экспонентным, а параметр — его натуральным параметром. Самое сложное в процедуре оценивания — интерпретация  $\mathbf{t}^*$ , т.е. вычисление значения  $\mathbf{s}^*$  канонического параметра, отвечающего  $\mathbf{t}^*$ , поскольку плотность вероятностного закона задана нам формулой именно через  $\mathbf{s}$ . Но эту трудность нельзя считать принципиальной, поскольку решать систему уравнений  $\text{grad } \Psi(\mathbf{s}) = \mathbf{t}^*$  надо лишь один раз в конце обработки. В наши дни для этой цели можно эффективно использовать персональный компьютер с универсальной стандартной программой решения таких систем.

Плотность  $p_0(\omega)$  определена с точностью до постоянного множителя, а направляющие статистики  $q_j(\omega)$  — с точностью до постоянного слагаемого. Если же принять  $p_0(\omega) = p_0(\omega; \mathbf{0})$ ,  $\mathbf{q}(\omega) = \mathbf{q}(\omega; \mathbf{0})$ , где  $\int \mathbf{q}(\omega; \mathbf{0}) p(\omega; \mathbf{0}) d\mu = \mathbf{0}$ , то  $\Psi(\mathbf{s}) = \mathcal{E}(P_s, P_0)$ , где относительная энтропия определена формулой (8), и

$$p(\omega; \mathbf{s}) = p(\omega; \mathbf{0}) \exp[s^j q_j(\omega; \mathbf{0}) - \mathcal{E}(P_s, P_0)]. \quad (17)$$

Сравним это представление (17) и представление плотности семейства гауссовских законов с единичной матрицей ковариаций и переменным средним  $\mathbf{s}$ :

$$p(x, \mathbf{s}) = p(x, \mathbf{0}) \exp[s^j x_j - (\mathbf{s}, \mathbf{s})/2]. \quad (18)$$

Из этого сравнения видно, во-первых (см. [2]), что гауссовские семейства являются единственными, в которых канонические параметризации совпадают с натуральными (для примера, в статистической физике натуральный параметр — температура  $t = s^{-1}$ ). И во-вторых, что относительная энтропия является обобщением половины квадрата евклидова расстояния. Аналогия заходит столь далеко, что для относительной энтропии справедлива (см. [21]) несимметричная теорема Пифагора.

**ТЕОРЕМА 4.** Если  $P_\sigma = \arg \min_{\mathbf{s}} \mathcal{E}(R, P_{\mathbf{s}})$ , где  $\gamma = \{P_{\mathbf{s}}\}$  — экспонентное семейство,  $\sigma \in \text{Int } \text{Dom } \gamma$ , то

$$\mathcal{E}(R, P_{\mathbf{s}}) = \mathcal{E}(R, P_\sigma) + \mathcal{E}(P_\sigma, P_{\mathbf{s}}) \quad \forall \mathbf{s} \in \text{Dom } \gamma. \quad (19)$$

Когда направляющие статистики  $q_j(\omega; \mathbf{0})$  в (17) ограничены, дополнительное условие на  $\sigma$  излишне.

В свое время Лаплас предложил измерять потери статистика расстоянием от сделанной им оценки до истинного значения аргумента. Но, как заметил

Гаусс, теория сильно упрощается, если считать потери  $\mathcal{L}$  квадратичной функцией ошибки оценивания; см. [22]. Для негауссовских вероятностных законов, как мы видим, естественно брать  $\mathcal{L} = 2\mathcal{E}(P^*, P_\theta)$ , а не какие-либо  $L_2(\mu)$ -нормы, зависящие от выбора  $\mu$ . Дальнейшая цель теоретика — указывать в задаче с.т.о. разумные априорные семейства с гарантируемой скоростью убывания характеристики погрешности (максимума риска, байесовского риска и т.п.). Отметим, что в отличие от  $L_1$ -подхода задача с.т.о. с энтропийной функцией потерь для всей доминированной подсовокупности  $\text{Card}(E, \mathcal{A}^*, Z)$  уже некорректна (см. [3, теорему 6]).

Выше мы уже упоминали, что задача статистического оценивания плотности по наблюдениям в самой широкой постановке не является равномерно состоятельной [3]. Построить равномерно сходящиеся алгоритмы ее восстановления удастся только для более узких семейств  $\mathcal{P}$  априори возможных законов. Проще всего предположить, что

$$C^{-1} \leq \frac{dQ}{dP}(\omega) \leq C \quad \forall \omega \in \Omega, \quad \forall P, Q \in \mathcal{P} \quad (20)$$

при некоторой положительной константе  $C$ . Заметим, что это условие является весьма сильным. В частности, все нормы  $L_p(R)$   $1 \leq p < \infty$ ,  $R \in \mathcal{P}$ , задают на  $\mathcal{P}$  одну и ту же топологию, все нормы  $L_2(R)$  эквивалентны при  $R \in \mathcal{P}$ . В таких условиях квазиинвариантности мер удается указывать нижние границы для точности восстановления плотности и строить алгоритмы, оптимальные по порядку точности [2, 23]. Однако класс семейств, удовлетворяющих (20), слишком узок. К нему не принадлежит, например, семейство (18) гауссовских законов.

Если же вместо  $L_2(R)$ -нормы разности в качестве функции потерь принять энтропию, для которой, как известно, всегда

$$\mathcal{E}(P, Q) \leq \|P - Q\|_P^2, \quad (21)$$

то разумным оказывается следующее достаточное условие квазиоднородности семейства  $\mathcal{P}$ , предложенное нами еще в [2]:

$$C^{-1} \leq \frac{D_{R'} \left[ \ln \frac{dP}{dQ}(\omega) \right]}{D_{R''} \left[ \ln \frac{dP}{dQ}(\omega) \right]} \leq C \quad (22)$$

при некоторой фиксированной константе  $C$  и любых  $P, Q, R', R'' \in \mathcal{P}$ . Этому условию локально удовлетворяют, например, семейства, зависящие от векторного параметра, регулярно гладкие в смысле Крамера–Рао. Наметим для них сейчас, следуя [2], основные результаты теории оптимального оценивания.

**ТЕОРЕМА 5.** Пусть  $\{P_\theta, \theta \in \Theta\}$  — компактное гладкое семейство вероятностных мер,  $\Theta$  — замкнутая область в  $R^n$ . Тогда при  $\mathcal{L}(\theta, P^*) = 2\mathcal{E}(P^*, P)$

$$\lim_{N \rightarrow \infty} N \cdot \inf_{\mathcal{H}(N)} \sup_{\theta \in \Theta} \mathcal{R}_{\mathcal{H}(N)}(\theta) = \dim \Theta, \quad (23)$$

где не обязательно ограничиваться оценками  $P^* \in \{P_\theta\}$ , а асимптотически оптимальной является оценка максимума правдоподобия  $\Pi$

$$\sup_{\theta \in \Theta} |N \cdot \mathcal{R}_{\Pi(N)}(\theta) - \dim \Theta| \xrightarrow{N \rightarrow \infty} 0. \quad (24)$$

Впрочем, и саму оценку максимума правдоподобия можно истолковать как оценку минимума относительной энтропии; см. [24, 25], а также [2].

Комбинируя теорему 5 с теоремой 4, мы развили в [2] для квазиоднородных семейств с правильно убывающими информационными поперечниками теорию почти оптимального “непараметрического” (точнее, счетно-параметрического) оценивания плотности. Как мы видели, этот наш подход близок к  $L_2$ -подходу, даже совпадает с ним на этапе построения средних значений от базисных (направляющих) статистик  $q_j(\omega)$ , но отличается в “интерпретации” построенных средних (и только при построении гистограммы оба подхода дают совпадающие результаты). Заинтересовавшийся читатель может ознакомиться с деталями по монографии [2]. Нам кажется, что благодаря своей эквивариантности экспонентный метод оценивания плотности не менее перспективен, чем ставшие традиционными; см, например, монографию [26], метод сглаживающего ядра или метод ортогонального разложения, рассмотренные в книге Девроя и Дьёрфи.

#### СПИСОК ЛИТЕРАТУРЫ

1. *Wald A.* Contributions to the theory of statistical estimation and testing hypotheses // *Ann. Math. Statist.* 1939. V. 10. № 4. P. 299–326.
2. *Ченцов Н.Н.* Статистические решающие правила и оптимальные выводы. — М.: Наука, 1972.
3. *Ченцов Н.Н.* О корректности задачи статистического точечного оценивания // *Теория вероятн. и ее примен.* 1981. Т. 26, вып. 1. С. 15–31; *Препр. Ин-та прикл. матем. им. М.В. Келдыша АН СССР.* — 1979. — 193.
4. *Гливиенко В.И.* Sulla determinazione empirica di una legge di distribuzione // *Giorn. dell'Istituto Italiano degli Attuari.* 1933. V. 4, № 1. P. 83–91.
5. *Соболь И.М.* Многомерные квадратурные формулы и функции Хаара. — М.: Наука, 1969.
6. *Abou-Jaoude S.* Conditions nécessaires et suffisantes de convergence  $L_1$  en probabilité de l'histogramme pour une densité // *Ann. de l'Inst. H. Poincaré. Ser. B.* 1976. V. 12. P. 213–231.
7. *Надарая Э.А.* О непараметрической оценке байесовского риска в задаче классификации // *Сообщ. АН ГрузССР.* 1976. Т. 82, № 2. С. 277–280.
8. *Ченцов Н.Н.* Категории математической статистики // *Докл. АН СССР.* 1965. Т. 164, № 3. С. 511–514.
9. *Morse N., Sacksteder R.* Statistical isomorphism // *Ann. Math. Statist.* 1966. V. 37, № 1. P. 203–214. [Рус. пер.: *Морзе Н., Сакстедер Р.* Статистический изоморфизм / *Математика.* 1968. Т. 12, № 6. С. 147–160.]
10. *Čencov N.N.* Algebraic foundations of mathematical statistics // *Math. Operationsforsch. Statist. Ser. Statistics.* 1978. V. 9, № 2. P. 267–276.
11. *Čencov N.N.* On basic concepts of mathematical statistics // *Banach Center Publ.* 1980. V. 6. P. 85–94.
12. *Csiszar I.* Information-type indices of the divergence of distributions. I, II // *Magyar Tud. Akad. Mat.* 1967. V. 17, № 2. P. 123–149. № 3. P. 267–291.
13. *Csiszar I.* On topological properties of  $f$ -divergence // *Studia Sci. Math. Hungar.* 1967. V. 2, № 3–4. P. 329–339.

14. *Morozova E.A., Cencov N.N.* Markov maps in noncommutative probability theory and mathematical statistics // Proc. 4th Intern. Vilnius Conf. Probability Theory Mathem. Statistics (1985). V. 2. — VNU Science Press, 1987. — P. 287–310.
15. *Ченцов Н.Н.* Оценка неизвестной плотности распределения по наблюдениям // Докл. АН СССР. 1962. Т. 147, № 1. С. 45–48.
16. *Фролов А.С., Ченцов Н.Н.* Использование зависимых испытаний в методе Монте-Карло для получения гладких кривых/ Тр. 6-го Всесоюзн. совещ. по теории вероятн. и матем. статистике (Вильнюс, 1960). — Вильнюс, 1962. — С. 425–437.
17. *Статулявичюс В.В.* Некоторые свойства асимптотически минимаксных статистических оценок плотности распределения: Дис. . . канд. физ.-мат. наук. — Вильнюс: Вильнюсский гос. университет, 1986. — 109 с.
18. *Стратонович Р.Л.* Быстрота сходимости алгоритмов оценки плотности распределения вероятностей // Изв. АН СССР. Сер. техн. киберн. 1969. Т. 6, № 1. С. 3–15.
19. *Ченцов Н.Н.* К систематической теории экспоненциальных семейств распределений вероятностей // Теория вероятн. и ее примен. 1966. Т. 11, вып. 3. С. 483–494.
20. *Barndorff-Nielsen O.* Information and Exponential Families in Statistical theory. — Wiley, 1978
21. *Ченцов Н.Н.* Несимметричное расстояние между распределениями вероятностей, энтропия и теорема Пифагора // Матем. заметки. 1968. Т. 4, вып. 3. С. 323–332.
22. *Le Cam L.* On some asymptotic properties of maximum likelihood estimates and related Bayes estimates // Univ. Calif. Publ. Statistics. 1953. V. 1. P. 267–329. [Рус. пер.: *Ле Кам Л.* О некоторых асимптотических свойствах оценок максимального правдоподобия и относительных байесовских оценках / Математика. — 1960. Т. 4, № 2. С. 69–119.]
23. *Ибрагимов И.А., Хасьминский Р.З.* Об оценке плотности распределения/ Сб. Исследования по математической статистике. IV // Записки научн. семинаров ЛОМИ. 1980. Т. 98. С. 61–86.
24. *Hartigan J.A.* The likelihood and invariance principles // Ann. Math. Statist. 1967. Т. 38. P. 533–539.
25. *Kriz T.A., Talacko J.V.* Equivalence of the maximum likelihood estimator to a minimum entropy estimator // Trab. Estadist. Invest. Oper. 1968. V. 19, № 1–2. P. 55–65.
26. *Надараля Э.А.* Непараметрическое оценивание плотности вероятностей и кривой регрессии. — Тбилиси: ТГУ, 1983.

## РАЗДЕЛ IV

# ГЕОМЕТРОСТАТИСТИКА И НЕКОММУТАТИВНАЯ ТЕОРИЯ ВЕРОЯТНОСТЕЙ

---

## ГЕОМЕТРИЯ “МНОГООБРАЗИЯ” РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ<sup>\*)</sup>

*Представлено академиком А.Н. Колмогоровым 9 IV 1964*

1. Пусть  $\Omega$  — произвольное множество элементарных исходов  $\omega$  с  $\sigma$ -алгеброй  $S$  измеримых подмножеств. Пусть, далее,  $I$  — идеал алгебры  $S$ . Рассмотрим максимальное “многообразие”  $H$  распределений вероятностей  $P\{d\omega\}$ , для каждого из которых идеал  $I$  есть класс всех событий нулевой вероятности. Наша цель — ввести в  $H$  “аффинную связность” и изучить “геометрию”  $H$ . Когда  $\Omega$  состоит из конечного числа атомов,  $H$  является многообразием в точном смысле этого слова. Вообще же говоря, изучаемая геометрия бесконечномерна.

2. Включим  $H$  в пространство  $W$  всех мер  $\sigma$ -конечной вариации, обязательно аннулирующейся на  $I$ , с конусом  $C$  строго положительных мер, аннулирующихся только на  $I$ . Через  $V \subset W$  обозначим пространство мер конечной вариации с положительным конусом  $K = V \cap C$ . Операция дифференцирования Радона–Никодима при фиксированной  $\mu \in C$  сопоставляет каждой мере  $\nu \in W$  функцию точки — ее производную

$$\nu\{\cdot\} \rightleftharpoons \frac{d\nu}{d\mu}(\cdot), \quad (1)$$

определенную с точностью до значений на множестве  $\mu$ -меры нуль; см. [1]. Если меры назвать векторами, а классы эквивалентных по  $\text{mod } I$  функций — ковекторами (элементами пространства  $W^*$ ), то билинейную операцию

$$\langle f, \nu \rangle = \int_{\Omega} f(\omega) \nu\{d\omega\}, \quad (2)$$

становящуюся на  $H$  определением математического ожидания, можно назвать

---

<sup>\*)</sup> Докл. АН СССР. 1964. Т. 158, № 3. С. 543–546.

“внутренним произведением” ковектора  $f$  и вектора  $\nu$ . “Многообразию”  $H$  выделяется из конусов  $C$  и  $K$  условием

$$\langle 1, P \rangle = 1. \quad (3)$$

С “поверхностью”  $H$  в каждой “точке”  $P$  связаны пространства  $L(P) = \{f : \langle |f|, P \rangle \text{ конечно}\}$ ,  $N(P) = \{f : \langle f, P \rangle = 0\}$ .

Касательное пространство  $\Lambda(P)$  к  $H$  в “точке”  $P$ , порожденное разностями  $P_\alpha - P$ , не зависит от выбора  $P$ ,  $\Lambda = \{R : \langle 1, R \rangle = 0\}$ .

**3.** Поскольку линейный оператор (1) ставит в соответствие каждому вектору свой ковектор и обратно, естественно сказать, что он определяет дважды ковариантный тензор. Далее, мы можем рассмотреть две квадратичные формы

$$(Q, R)_P = \int \frac{dQ}{dP}(\omega) \frac{dR}{dP}(\omega) P\{d\omega\}, \quad (4)$$

$$(f, g)^P = \int f(\omega) g(\omega) P\{d\omega\}, \quad (5)$$

задающие скалярные произведения, связанные с точкой  $P$ . Кажется соблазнительным (см. [2, 3]) объявить  $H$  “римановым пространством с “метрикой” (4). Однако такой подход не улавливает всех “статистических” свойств семейств распределений вероятностей. Пока мы лишь воспользуемся изоморфизмом

$$W \sim W^*, \quad C \sim C^*, \quad V \sim L(P), \quad \Lambda \sim N(P), \quad P \sim 1, \quad (6)$$

устанавливаемым оператором (1), чтобы декретировать прямую  $\lambda \cdot P\{\cdot\}$  нормалью к “поверхности”  $H$  в точке  $P$ .

**4.** Каждому параллельному переносу пространства  $W^* : f(\cdot) \xrightarrow{g} f(\cdot) + g(\cdot)$  соответствует автоморфизм конуса  $C$

$$\mu\{\cdot\} \xrightarrow{g} \nu\{\cdot\}, \quad \text{где} \quad \frac{d\nu}{d\mu}(\omega) = \exp(g(\omega)). \quad (7)$$

Коммутативная группа  $G$  всех “трансляций” конуса  $C$  просто транзитивна, ибо  $\mu\{\cdot\} \xrightarrow{g} \nu\{\cdot\}$  тогда и только тогда, когда  $g(\omega) = \ln \frac{d\nu}{d\mu}(\omega)$ .

Конус  $K$  не инвариантен относительно  $G$ . Но поскольку для каждого  $g$  не пусто  $K \cap g(K)$ , то  $G$  является псевдогруппой (обобщенных) “автоморфизмов”  $K$  (см. [4]);  $H$  — “гиперповерхность” в  $K$ . “Автоморфизм” касательного пространства  $\Lambda$ , индуцируемый в смысле Леви-Чивита (см. [5]) трансляцией  $P_1\{\cdot\} \rightarrow P_2\{\cdot\}$ , записывается формулой

$$dR_1 \rightarrow dR_2 = \left( \frac{dR_1}{dP_1} - c \right) dP_2, \quad (8)$$

$$R_2(B) = \int_B \frac{R_1\{d\omega\} P_2\{d\omega\}}{P_1\{d\omega\}} - P_2(B) \int_\Omega \frac{R_1\{d\omega\} P_2\{d\omega\}}{P_1\{d\omega\}}.$$

Если в (8) перейти к плотностям, то формула упрощается:

$$\frac{dR_1}{dP_1} \rightarrow \frac{dR_1}{dP_1} - c = \frac{dR_2}{dP_2}. \quad (9)$$

Сопряженное сохраняющее “внутреннее произведение” преобразование двойственного подпространства  $N(P)$  записывается

$$f_1(\omega) \rightarrow f_2(\omega) = f_1(\omega) \frac{dP_1}{dP_2}(\omega).$$

Транзитивная псевдогруппа “автоморфизмов” (7) порождает расслоенное пространство с базой  $H$  и слоями  $\Lambda(P)$ . В индуцируемой в  $H$  этим расслоением “аффинной связности” параллелизм векторов абсолютен (см. [5]), так что связность является плоской и даже локально аффинной (см. [6]).

5. Фиксируем произвольную меру  $\mu \in H$  и будем писать  $p(\omega) = \frac{dP}{d\mu}(\omega)$ . Условие параллельности переноса касательного вектора вдоль геодезической линии  $p(\omega; t)$  в силу (9) записывается

$$\lambda(t) \frac{p'_t(\omega; t)}{p(\omega; t)} = \frac{p'_t(\omega; 0)}{p(\omega; 0)} - \psi(t). \quad (10)$$

Известной заменой переменного (см. [5, 6]) можно перейти к каноническому аффинному параметру  $s$ :

$$q(\omega; s) = \frac{\partial}{\partial s} \ln p(\omega; s) = q(\omega; 0) - \varphi(s), \quad (11)$$

где  $s$  определено с точностью до линейной подстановки  $s \rightarrow as + b$ . Из “первого интеграла” (11) нетрудно получить явное выражение для “геодезического семейства”  $p(\omega; s)$ . Воспользуемся, однако, тем, что геометрия  $H$  есть “проективная геометрия” пучка прямых конуса  $C$  с трансляциями (7), поскольку прямой выкладкой проверяется, что “проективное” определение параллельного перенесения бесконечно малых касательных векторов совпадает с (8). Геодезические в  $C$  траектории однопараметрических подгрупп группы  $G$  имеют вид

$$f(\omega; s) = f(\omega; 0) \exp[s \cdot g(\omega)]. \quad (12)$$

Геодезические в  $H$  обязаны быть “логарифмическими проекциями” на  $H$  этих траекторий, т.е. отличаться от  $f(\omega; s)$  нормирующим множителем. Таким образом, геодезическая, проходящая через данную точку  $p(\omega; 0)$  в данном направлении, описывается формулой

$$p(\omega; s) = \frac{1}{a(s)} p(\omega; 0) \exp[s \cdot q(\omega; 0)], \quad (13)$$

где  $a(s)$  — нормирующая константа

$$a(s) = M_0 \exp[s \cdot q(\omega; 0)] = \int \exp[s \cdot q(\omega; 0)] p(\omega; 0) d\mu, \quad (14)$$

связанная с  $\varphi(s)$  соотношением

$$\varphi(s) = \frac{d}{ds} \ln a(s) = M_s q(\omega; 0). \quad (15)$$

Геодезическая, проходящая через две точки, описывается в барицентрических аффинных координатах формулой

$$p(\omega; \tau) = \frac{1}{b(\tau)} [p(\omega; 0)]^{1-\tau} [p(\omega; 1)]^\tau, \quad (16)$$

где  $b(\tau)$  — опять-таки нормирующая константа

$$b(\tau) = \int [p(\omega; 0)]^{1-\tau} [p(\omega; 1)]^\tau d\mu. \quad (17)$$

**6.** При некоторых значениях  $s$  траектория может выйти из конуса  $K$ , и соответствующие плотности (и меры) нельзя будет отнормировать. Нормирующая константа  $A(s)$  траектории (12) совпадает с производящей функцией моментов распределения значений функции  $g(\omega)$  относительно меры с плотностью  $f(\omega; 0)$  (см. [7])<sup>1</sup>. Справедливо следующее предложение.

*Если плотности  $f_0(\omega)$  и  $f_1(\omega)$  соответствуют мерам  $\nu_0 \in K$  и  $\nu_1 \in K$ , то все точки  $\nu_z$  (с плотностями  $f_0^{1-z} \cdot f_1^z$ ) отрезка  $0 \leq z \leq 1$  (области  $0 \leq \operatorname{Re} z \leq 1$ ) траектории, проходящей через  $\nu_0$  и  $\nu_1$ , также принадлежат  $K$ . Нормирующая константа  $A(z) = \langle 1, \nu_z \rangle$  является аналитической функцией  $z$  в полосе  $0 < \operatorname{Re} z < 1$  и непрерывной при  $0 \leq \operatorname{Re} z \leq 1$ .*

Простым следствием отсюда является теорема:

*“Многообразие”  $H$  геодезически выпукло.*

В силу выпуклости геодезическая в  $H$ , соответствующая данной траектории в  $C$ , может быть: а) всей аффинной прямой; б) лучом; в) отрезком; г) точкой; д) пустым множеством (причем граничные значения в случаях б) и в) могут исключаться). Внутри области существования геодезическая является аналитической линией. В граничных точках справедливые внутри аналитические формулы могут терять смысл.

**7.** Величина  $a(s)$  сначала монотонно убывает до  $a(0) = 1$ , затем монотонно возрастает, не медленнее экспоненты;

$$a^{(n)}(s) = a(s) M_s [q(\omega; 0)]^n = a(s) \int [q(\omega; 0)]^n p(\omega; s) d\mu.$$

При  $s = 0$   $a'(0) = 0 = \varphi(0)$ ,  $a''(0) = \varphi'(0) = D_0 q(\omega; 0) = I_F$ ; последняя величина носит название информационного количества Фишера [2, 8]. Сходным образом изменяется величина  $b(\tau)$ :

$$b(\tau) = \int \left[ \frac{dP_1}{dP_0}(\omega) \right]^\tau P_0 \{d\omega\} = \int \left[ \frac{dP_0}{dP_1}(\omega) \right]^{1-\tau} P_1 \{d\omega\}.$$

Сначала  $b(\tau)$  монотонно убывает до некоторого минимального значения  $b(\tau_0) = \rho(P_0, P_1) < 1$ ,  $0 < \tau_0 < 1$ , затем монотонно возрастает. В [7] величину  $I_D = -\ln \rho$  было предложено применять как меру различия между распределениями

<sup>1</sup>Если  $f(\omega; 0)$  — плотность распределения вероятностей случайных исходов  $\omega$ ,  $A(is)$  — характеристическая функция случайной величины  $g(\omega)$ .

(гипотезами)  $P_0$  и  $P_1$ . Если обозначить  $l(\omega) = \ln p(\omega; 1) - \ln p(\omega; 0)$ , то  $b^{(n)}(\tau) = b(\tau) M_\tau [l(\omega)]^n$ . Величину  $-b'(0) = -M_0 l(\omega) = I_C$  называют шенноновским количеством информации об отличии  $P_0$  от  $P_1$  [8]. Это понятие обобщено в [9], где были введены количества информации порядка  $\tau$ :  $I_R(\tau) = (1-b(\tau))/\tau$ ,  $I_R(0) = I_C$ . Как известно (см. [8]), при соответствующем предельном переходе  $I_C \rightarrow I_F$ , когда  $P_1 \rightarrow P_0$  вдоль геодезической. Формулы (13) и (16) связаны соотношениями

$$q(\omega; 0) = l(\omega) - M_0 l(\omega), \quad \varphi(\tau) = M_\tau l(\omega) - M_0 l(\omega).$$

Для  $n$ -го семиинварианта  $\varkappa_s^n q(\omega)$  распределения значений функции  $q(\omega)$  по распределению исходов  $\omega$  с плотностью  $p(\omega; s)$

$$\varphi^{(k)}(s) = \varkappa_s^k q(\omega; 0) = \varkappa_s^k q(\omega; s), \quad k \geq 1, \quad (18)$$

откуда каноническое уравнение геодезической имеет вид

$$\frac{\partial q(\omega; s)}{\partial s} + D_s q(\omega; s) = \frac{\partial q(\omega; s)}{\partial s} + \int [q(\omega; s)]^2 p(\omega; s) d\mu = 0. \quad (19)$$

**8.** В силу локальной аффинности  $H$  расслаивается на вполне геодезические поверхности. Например, геодезическая поверхность, проходящая через три точки, описывается формулой

$$p(\omega; \alpha, \beta, \gamma) = \frac{1}{b(\alpha, \beta, \gamma)} [p_1(\omega)]^\alpha [p_2(\omega)]^\beta [p_3(\omega)]^\gamma, \quad \alpha + \beta + \gamma = 1. \quad (20)$$

Любая геодезическая, проходящая через две “точки” этого семейства, также принадлежит ему. Двумерный аналог формулы (13) имеет вид  $\ln p(\omega; s, t) = -\ln a(s, t) + tq_t(\omega; 0, 0) + sq_s(\omega; 0, 0)$ .

Геометрическую характеристику ранга  $r$  семейства взаимно абсолютно непрерывных распределений вероятностей  $p(\omega; \theta)$  (см. [10]) дает теорема:

*Если семейство  $p(\omega; \theta)$  имеет ранг  $r$ , то выпуклая геодезическая оболочка семейства есть область некоторой  $r$ -мерной геодезической поверхности, и обратно.*

**9.** В  $H$  существуют метрические связности, совместимые с аффинной связностью (8). Достаточно задать дифференциальную квадратичную форму в одной точке  $\mu$  и затем перенести ее параллельно во все остальные. Например, если в точке метрический тензор есть тензор Радона–Никодима (1), то

$$(Q, R)_P = \text{Cov}_\mu \left( \frac{dQ}{dP}, \frac{dR}{dP} \right) = \int \left[ \frac{dQ}{dP} - \int \frac{dQ}{dP} d\mu \right] \left[ \frac{dR}{dP} - \int \frac{dR}{dP} d\mu \right] d\mu. \quad (21)$$

Отсюда вытекает, в частности, что тензор (1) не является стационарным. Длина кривой в любой из таких метрик является изменением соответствующего аффинного параметра.

**10.** Многие свойства замечательных семейств распределений вероятностей связаны с тем, что они являются геодезическими семействами. К числу таковых относятся распределения Гаусса и Пуассона. Интересно отметить, что все распределения Пирсона [11], заданные уравнением

$$y'(x) = t \frac{x - \alpha}{(x - \beta)(x - \gamma)} y(x),$$

естественно распадаются на геодезические  $y(x; t)$ . Объем заметки не позволяет коснуться других вопросов, где геометрический подход также оказывается полезным.

В заключение автор считает приятным долгом выразить благодарность Е.А. Морозовой за ценную дискуссию.

#### СПИСОК ЛИТЕРАТУРЫ

1. Халмош П. Теория меры. — М.: ИЛ, 1953.
2. Kullback S. // Ann. Math. Stat. 1954. V. 25, № 4. P. 745.
3. Ченцов Н.Н. // Докл. АН СССР. 1962. Т. 147, № 1. С. 45.
4. Ehresmann S., Chern S. S. Géométrie différentielle. — Paris: Colloq. intern. du C.N.R.S., 1953.
5. Норден А.П. Пространства аффинной связности. — М.-Л., 1950.
6. Фавар Ж. Курс локальной дифференциальной геометрии. — М.: ИЛ, 1960.
7. Chernoff H. // Ann. Math. Stat. 1952. V. 23, № 4. P. 493.
8. Kullback S., Leibler R.A. // Ann. Math. Stat. 1951. V. 22, № 1. P. 79.
9. Rényi A. // Proc. 4-th Berkeley Simpos. Math. Stat. and Prob. № 1. 1961. P. 547.
10. Дынкин Е.Б. // Успехи матем. наук. 1951. Т. 6, вып. 1. С. 66.
11. Крамер Г. Математические методы статистики. — М.: ИЛ, 1948.

## КАТЕГОРИИ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ \*)

Представлено академиком А.Н. Колмогоровым 25 II 1965

Условимся обозначать  $\text{Car}(\Omega, S)$  совокупность всех распределений вероятностей на измеримом пространстве  $(\Omega, S)$ , т.е. совокупность всех нормированных мер на  $\sigma$ -алгебре  $S$  измеримых подмножеств пространства  $\Omega$  элементарных исходов  $\omega$ . Для любой пары объектов  $\text{Car}(\Omega_1, S_1)$  и  $\text{Car}(\Omega_2, S_2)$  существуют марковские (гомо-) морфизмы  $\Pi: \text{Car}_1 \rightarrow \text{Car}_2$ , ставящие в соответствие каждому распределению  $P \in \text{Car}_1$  свое распределение  $Q \in \text{Car}_2$  по формуле

$$Q\{\cdot\} = \int_{\Omega_1} \Pi\{\cdot|\omega'\} P\{d\omega'\}. \quad (1)$$

Переходная мера  $\Pi\{A|\omega'\}$  в (1) —  $S_1$ -измеримая функция на  $\Omega_1$  при фиксированном  $A \in S_2$  и нормированная мера (распределение вероятностей) на  $S_2$  при фиксированном  $\omega' \in \Omega_1$ .

Марковские морфизмы образуют мультипликативную систему:

Cat 1. Для каждого объекта  $\text{Car}(\Omega, S)$  существует и единствен тождественный автоморфизм. Он задается переходной мерой  $\Pi\{A|\omega\} = \chi_A(\omega)$  — индикатором переменного множества  $A$ .

\*) Докл. АН СССР. 1965. Т. 164, № 3. С. 511–514.

Cat 2. Композиция марковских морфизмов  $\Pi_{12}(\text{Cat}_1 \rightarrow \text{Cat}_2)$  и  $\Pi_{23}(\text{Cat}_2 \rightarrow \text{Cat}_3)$  определена и является, в силу теоремы Фубини, марковским морфизмом  $\Pi_{13}(\text{Cat}_1 \rightarrow \text{Cat}_3)$  с переходной мерой

$$\Pi_{13}\{A|\omega'\} = \int_{\Omega_2} \Pi_{12}\{d\omega''|\omega'\} \Pi_{23}\{A|\omega''\}.$$

Таким образом, объекты  $\text{Cat}(\Omega, S)$  с системой марковских морфизмов образуют категорию (см. [1]) — категорию  $\text{CAP}$  совокупностей распределений вероятностей.

Объектами изучения математической статистики являются семейства  $w = \{w_\theta\}$  распределений вероятностей  $P_\theta\{d\omega\}$  на одном и том же измеримом пространстве  $(\Omega, S)$ , т.е. всевозможные подмножества совокупностей  $\text{Cat}(\Omega, S)$ . Получается следующая “геометрическая картина”. Имеются “пространства”  $\text{Cat}(\Omega, S)$  и система отображений. Изучаются свойства “фигуры”  $w$ , инвариантные относительно “обратимых” отображений этой фигуры. Возникающие отношения можно описать на языке категорий, поскольку семейства  $w$  сами образуют категорию с системой марковских морфизмов<sup>1)</sup>, правильнее сказать, даже две категории  $\text{FAM}_1$  и  $\text{FAM}_0$ , в зависимости от того, включается способ параметризации в определение семейства  $w$  или нет. В категории  $\text{FAM}$  обычное отношение эквивалентности объектов конкретизируется в следующее.

**ОПРЕДЕЛЕНИЕ 1.** Два семейства  $w_1$  и  $w_2$  вероятностных пространств  $w_\theta^{(1)} = (\Omega_1, S_1, P_\theta^{(1)})$ ,  $w_\theta^{(2)} = (\Omega_2, S_2, P_\theta^{(2)})$  называются *эквивалентными*, если они одинаково запаараметризованы и существуют переходные меры  $\Pi_{12}\{d\omega''|\omega'\}$  и  $\Pi_{21}\{d\omega'|\omega''\}$  такие, что  $w_1\Pi_{12} = w_2$  и  $w_2\Pi_{21} = w_1$  ( $w_\theta^{(1)}\Pi_{12} = w_\theta^{(2)}$ ,  $w_\theta^{(2)}\Pi_{21} = w_\theta^{(1)}$  при всех  $\theta$ ).

Покажем, что описанный выше геометрический подход в стиле Кляйна естественно возникает в теории статистических решений [2], а определение 1 описывает семейства распределений с “одинаковыми” статистическими свойствами. Пусть, кроме семейства  $w = \{w_\theta\}$  распределений вероятностей  $P_\theta\{d\omega\}$  исходных  $\omega$  наблюдений, задано измеримое пространство  $(\mathcal{E}, B)$  выводов  $\varepsilon$  и указан некоторый стохастический закон  $\Pi$  вывода — распределение вероятностей  $\Pi\{d\varepsilon|\omega\}$  выводов  $\varepsilon$  при известном исходе  $\omega$ ; см. [3]. Каждое правило  $\Pi$  индуцирует как раз по формуле (1) семейство

$$Q_\theta^{(\Pi)}\{\cdot\} = \int_{\Omega} \Pi\{\cdot|\omega\} P_\theta\{d\omega\} \quad (2)$$

распределений вероятностей выводов, или, для краткости, некоторое статистическое решение проблемы  $\mathcal{E}$  для семейства  $w$ . Все семейства  $Q_\theta \subset \text{Cat}(\mathcal{E}, B)$ , представимые в виде (2) при данном  $w$  и всевозможных  $\Pi$ , образуют класс  $D(\mathcal{E}, w)$  всех решений проблемы  $\mathcal{E}$  для  $w$ .

<sup>1)</sup>Превращение собрания объектов  $\text{FAM}$  в категорию достигается ценой усложнения мультипликативной системы морфизмов: каждый морфизм  $w_1 \rightarrow w_2$  индексируется интегральным ядром  $\Pi\{\cdot|\cdot\}$  и объектом  $w_1$ ; морфизмы  $w_1 \rightarrow w_2$ , определяющие одинаковые отображения элементов  $w_1$  в элементы  $w_2$ , отождествляются.

ПРИМЕР. Когда пространством выводов является множество  $\Theta$  всех значений параметра  $\theta$ , проблема  $\Theta$  для семейства  $\{\omega_\theta\}$  есть задача об оценке по наблюдаемому исходу  $\omega$  неизвестного параметра закона распределения  $P_\theta$  наблюдаемых исходов.

Если отвлечься от сложности и стоимости реализации процедуры вывода, то естественным будет

ОПРЕДЕЛЕНИЕ 2. Два семейства  $w_1$  и  $w_2$  статистически эквивалентны, когда для любой проблемы  $\mathcal{E}$  классы  $D(\mathcal{E}, w_1)$  и  $D(\mathcal{E}, w_2)$  одинаковы.

ТЕОРЕМА 1. Статистическая эквивалентность семейств распределений есть эквивалентность объектов категории FAM.

Установим ряд предложений “марковской геометрии”.

ТЕОРЕМА 2. Любые два простых семейства, т.е. состоящие каждое из единственного распределения, статистически эквивалентны.

Если  $w_1\Pi_{12} = w_2$  и  $w_2\Pi_{21} = w_1$ , то  $w_1\Pi_{12}\Pi_{21} = w_2$  и  $w_2\Pi_{21}\Pi_{12} = w_1$ , т.е. семейства  $w_1$  и  $w_2$  являются подмножествами стационарных “векторов” неотрицательных линейных преобразований  $\Pi_{12}\Pi_{21}$  и  $\Pi_{21}\Pi_{12}$ .

ЛЕММА 1. Пусть  $V_1$  и  $V_2$  — минералы всех мер ограниченной вариации, стационарных соответственно для марковских линейных операторов  $\Pi_{12}\Pi_{21}$  и  $\Pi_{21}\Pi_{12}$ , и  $v_1 = V_1 \cap \text{Car}_1$ ,  $v_2 = V_2 \cap \text{Car}_2$ . Тогда операторы  $\Pi_{12}$  и  $\Pi_{21}$  устанавливают взаимно однозначное линейное соответствие между “точками” минералов  $V_1$  и  $V_2$ , а также между “точками” семейств  $v_1$  и  $v_2$ .

Ограничимся пока совокупностями всех распределений вероятностей на множестве  $\Omega$ , имеющих лишь конечное число исходов и все подмножества которых измеримы. Такие совокупности образуют полную в CAR подкатегорию, которую мы обозначим CARF. Мультипликативная система марковских операторов категории CARF состоит из стохастических матриц. Строение множества стационарных распределений конечной цепи Маркова хорошо известно; см. [4]. Это позволяет дать полное описание эквивалентных семейств категории FAMF.

ТЕОРЕМА 3. Два семейства  $w_1 \subset \text{Carf}(\Omega_1)$  и  $w_2 \subset \text{Carf}(\Omega_2)$  распределений вероятностей эквивалентны тогда и только тогда, когда совпадают семейства факторов-распределений их минимальных достаточных статистик:

$$P_\theta^{(1)}\{\omega'\} = R_\theta\{\varepsilon(\omega')\} \cdot Q^{(1)}\{\omega'|\varepsilon\},$$

$$P_\theta^{(2)}\{\omega''\} = R_\theta\{\varepsilon(\omega'')\} \cdot Q^{(2)}\{\omega''|\varepsilon\};$$

здесь  $Q^{(i)}$  не зависит от  $\theta$ ,  $R_\theta$  не зависит от  $i$ .

Для всей категории FAM удается сформулировать лишь более слабое предложение.

ТЕОРЕМА 3'. Пусть семейство  $\{P_1, P_2\}$  эквивалентно  $\{Q_1, Q_2\}$ , причем  $Q_i = P_i\Pi_{12}$ ,  $P_j = Q_j\Pi_{21}$ . Обозначим

$$P_e = \frac{P_1 + P_2}{2}, \quad Q_e = \frac{Q_1 + Q_2}{2}, \quad p(\omega') = \frac{dP_1}{dP_e}(\omega'), \quad q(\omega'') = \frac{dQ_1}{dQ_e}(\omega'').$$

Тогда:

- 1)  $P_e p^{-1}\{\cdot\} = Q_e q^{-1}\{\cdot\}$ ;  
 2)  $P_f \Pi_{12} = Q_f$ ,  $Q_f \Pi_{21} = P_f$ , где  $\frac{dP_f}{dP_e}(\omega') = f(p(\omega'))$ ,  $\frac{dQ_f}{dQ_e}(\omega'') = f(q(\omega''))$ .

Аналогичные предложения справедливы для любых конечных семейств.

Рассмотрим теперь для конечных множеств  $\Omega$  конусы  $\text{Conf}(\Omega)$  всех неотрицательных мер на алгебре всех подмножеств  $\Omega$ . Неотрицательные линейные гомоморфизмы  $R: \text{Conf}_1 \rightarrow \text{Conf}_2$ ,

$$\nu\{\cdot\} = \int_{\Omega_1} R\{\cdot|\omega'\} \mu\{d\omega'\}, \quad (3)$$

уже с ненормированными переходными мерами задаются произвольными неотрицательными матрицами  $\|r_j^i\|$ . Совершенно аналогично предыдущему конусы  $\text{Conf}(\Omega)$  образуют категорию  $\text{CONF}$  с системой морфизмов вида (3). При этом для семейств распределений вероятностей как для “фигур” объекта  $\text{Conf}$  возникает более широкое отношение аффинной эквивалентности.

Для каждой меры из конуса  $\text{Conf}(\Omega)$  существует и единственна проекция на “гиперповерхность”  $\text{Carf}(\Omega)$

$$\pi\mu\{\cdot\} = \frac{1}{\mu\{\Omega\}} \mu\{\cdot\}. \quad (4)$$

Оператор проектирования  $\pi$  определяет ковариантный функтор, переводящий объекты категории  $\text{CONF}$  в объекты категории  $\pi\text{CARF}$  с мультипликативной системой положительных проективных морфизмов вида  $\pi R$ . Каждый проективный морфизм задается переходной мерой  $R\{A|\omega\}$ , т.е. неотрицательной матрицей  $\|r_j^i\|$ , определенной с точностью до положительного числового множителя.

У фиксированной совокупности  $\text{Carf}$  система ее марковских эндоморфизмов образует полугруппу  $\mathfrak{M}$  с единицей. Она содержит подгруппу  $\mathfrak{S}$  марковских автоморфизмов. Для  $\text{Carf}(\Omega)$  группа  $\mathfrak{S}$  состоит лишь из перенумераций исходов  $\omega_i$  множества  $\Omega$ . Для конуса  $\text{Conf}(\Omega)$  группа  $\mathfrak{P}$  положительных линейных автоморфизмов много богаче. Компонентой единицы в  $\mathfrak{P}$  является коммутативная просто транзитивная группа  $\mathfrak{L}$  положительных диагональных матриц, а факторгруппа  $\mathfrak{P}/\mathfrak{L}$  изоморфна  $\mathfrak{S}$ .

Конус  $\text{Conf}(\Omega)$  содержит конус  $C(\Omega)$  всех строго положительных мер на  $\Omega$ . Относительно группы  $\mathfrak{L}$   $C(\Omega)$  является аффинным пространством; см. [5]. Обозначим через  $H(\Omega)$  гиперповерхность  $C(\Omega) \cap \text{Carf}(\Omega)$ ;  $H(\Omega)$  инвариантна относительно группы  $\mathfrak{L}$  трансляций. Однако справедлива

**ТЕОРЕМА 4.** Если фигура  $w_1 \subset H(\Omega)$  непрерывно передвигается внутри  $H(\Omega)$  с сохранением эквивалентности в фигуру  $w_2 \subset H(\Omega)$  семейством марковских эндоморфизмов ( $w_1$   $\mathfrak{M}$ -гомотопна  $w_2$ ), то это передвижение можно осуществить семейством трансляций ( $w_1$   $\mathfrak{L}$ -гомотопна  $w_2$ ).

Максимальная группа проективных эндоморфизмов  $\text{Carf}$  есть факторгруппа группы  $\mathfrak{P}$  по группе  $\mathfrak{L}$  скалярных матриц. Компонента единицы в  $\mathfrak{P}/\mathfrak{L}$  есть просто транзитивная коммутативная группа  $\mathfrak{S} = \mathfrak{L}/\mathfrak{L}$ .

ТЕОРЕМА 5. Многообразие  $H$  с группой движений  $\mathfrak{G}$  является однородным локально аффинным пространством<sup>2)</sup>. Его естественная геометрия линейной связности есть внутренняя в смысле Леви–Чивита геометрия гиперповерхности  $H$ , оснащенной в  $S$  псевдонормальными вида  $OP$  в переменной точке  $P \in H$ .

Из теоремы 5 вытекает, что два  $\mathfrak{M}$ -гомотопных семейств  $w_1, w_2 \subset H(\Omega)$  обязательно  $\mathfrak{G}$ -конгруэнтны.

ТЕОРЕМА 6. Единственной линейной связностью многообразия  $H(\Omega)$ , симметричной относительно перенумераций исходов и сохраняющей равенство  $\mathfrak{M}$ -гомотопных касательных векторов, является его естественная связность.

Теорема 6 показывает значение введенной нами в [5] естественной линейной связности на  $H$  в вопросах математической статистики.

Вернемся снова к произвольным измеримым пространствам  $(\Omega, S)$ . Если отказаться в (1) от нормированности переходной меры и не предполагать ее равномерную по  $\omega'$  ограниченность, то в (3) приходится рассматривать конусы  $\text{Con}(\Omega, S)$  всех неотрицательных мер, как конечных, так и бесконечных, в том числе необязательно  $\sigma$ -конечных. Они образуют категорию  $\text{CON}$ . Определив на конечных мерах операцию проектирования по (4), мы зададим функтор, переводящий объекты категории  $\text{CON}$  в объекты категории  $\pi\text{CAP}$ . Соответственно возникает категория  $\pi\text{FAM}$  с системой морфизмов, более широкой, чем у категории  $\text{FAM}$ :

$$Q_\theta\{\cdot\} = \frac{1}{\lambda(P_\theta)} \int_{\Omega_1} R\{\cdot|\omega'\} P_\theta\{d\omega'\}, \quad \text{где } \lambda(P_\theta) = \int_{\Omega_1} R\{\Omega_2|\omega'\} P_\theta\{d\omega'\}. \quad (5)$$

Эта категория тесно связана с теорией стохастического моделирования (теорией метода Монте-Карло). Представим ядро из (5) в виде  $R\{\cdot|\omega'\} = g(\omega') \Pi\{\cdot|\omega'\}$ , где  $g(\omega') = R\{\Omega_2|\omega'\} P_\theta$ -почти всюду конечна,  $\Pi\{\cdot|\omega'\}$  — распределение вероятностей. Тогда для любой  $Q_\theta$ -интегрируемой функции  $f(\omega'')$

$$\left[ \frac{1}{N} \sum_{i=1}^N g(\omega'_i) f(\omega''_i) \right] \left[ \frac{1}{N} \sum_{i=1}^N g(\omega'_i) \right]^{-1} \implies \int_{\Omega_2} f(\omega'') Q_\theta\{d\omega''\},$$

где  $\omega'_i \in \Omega_1$  — независимые наблюдения с распределением  $P_\theta\{d\omega'\}$ , а  $\omega''_i$  — дополнительные наблюдения, разыгранные со случайными распределениями  $\Pi\{d\omega''|\omega'_i\}$ .

Можно продолжить параллель между категориями  $\text{FAM}$  и  $\pi\text{FAM}$ <sup>3)</sup>. Однако, в противоположность конечному случаю, связать таким путем с объектом  $\text{Cap}(\Omega, S)$  естественную линейную связность не удастся. Приходится с самого начала рассматривать более узкие конусы  $C(\Omega, S, I)$  строго положительных  $\sigma$ -конечных мер с общим идеалом  $I$  множеств меры нуль и соответственно более

<sup>2)</sup> У всякого однородного пространства отрицательной кривизны существует естественная граница; см. [6, 7]. Для пространств нулевой кривизны замыкания неединственны. Однако если вместо связной группы  $\mathfrak{L}/\Lambda$  рассмотреть полную группу  $\mathfrak{F}/\Lambda$  автоморфизмов и непрерывную подгруппу  $\mathfrak{M}$  марковских эндоморфизмов открытого симплекса  $H(\Omega)$ , то разумное замыкание  $H(\Omega)$  становится единственным. Соответствующая граница была впервые построена в [8], где  $H(\Omega)$  замыкали так, чтобы сохранялась непрерывность условных вероятностей.

<sup>3)</sup> В частности, истолковать возникающее здесь понятие эквивалентности, описать строение проективно эквивалентных семейств на множествах с конечным числом исходов и т.п.

узкие объекты  $H(\Omega, S, I) = \text{Cap}(\Omega, S) \cap C(\Omega, S, I)^4$ . Возникающая здесь естественная линейная связность была описана нами в [5].

### СПИСОК ЛИТЕРАТУРЫ

1. Гротендик А. О некоторых вопросах гомологической алгебры. — М.: ИЛ, 1961.
2. Вальд А. Последовательный анализ. — М., 1960.
3. Bahadur R.R. // Ann. Math. Statist. 1954. V. 25, № 3, P. 423.
4. Крейн М.Г., Рутман М.А. // Успехи матем. наук. 1948. Т. 3, вып. 1. С. 3.
5. Ченцов Н.Н. // Докл. АН СССР. 1964. Т. 158, № 3. С. 543.
6. Гельфанд И.М., Граев М.И. // Тр. Моск. матем. общ. 1959. Т. 8. С. 321.
7. Карпелевич Ф.И. Дис. ... д-ра физ.-мат. наук. — М., 1963.
8. Воробьев Н.Н., Фадеев Д.К. // Теория вероятн. и ее примен. 1961. Т. 6, вып. 1. С. 116.
9. Антоновский М.Я., Болтянский В.Г., Сарымсаков Т.А. Топологические алгебры Буля. — Ташкент, 1963.

## К СИСТЕМАТИЧЕСКОЙ ТЕОРИИ ЭКСПОНЕНЦИАЛЬНЫХ СЕМЕЙСТВ РАСПРЕДЕЛЕНИЙ ВЕРОЯТНОСТЕЙ<sup>\*)</sup>

1. Пусть  $\mu\{d\omega\}$  — некоторая конечная мера на измеримом пространстве  $(\Omega, S)$ . Рассмотрим многообразие  $\mathbf{H}(\Omega, S, I)$  — совокупность всех распределений вероятностей  $P\{d\omega\}$  на  $(\Omega, S)$ , взаимно абсолютно непрерывных относительно  $\mu\{d\omega\}$ , т.е. имеющих с  $\mu\{\cdot\}$  общий идеал  $I$  множеств меры нуль. Такая совокупность полностью определяется заданием идеала  $I$  и не зависит от конкретного выбора меры  $\mu$ . Однако с каждой такой мерой  $\mu$  связана, образно говоря, своя “система координат”, в которой всякое  $P\{d\omega\} \in \mathbf{H}$  описывается своей производной Радона–Никодима  $p(\omega) = \frac{dP}{d\mu}(\omega)$  —  $S$ -измеримой функцией, определенной с точностью до значений на нуль-множестве,  $\int_{\Omega} p(\omega) \mu\{d\omega\} = 1$ . При замене основной меры  $\mu$  на взаимно абсолютно с ней непрерывную  $\sigma$ -конечную меру  $\nu$  все плотности умножаются на постоянную функцию  $\frac{d\mu}{d\nu}(\omega)$ . Линейные функционалы на  $\mathbf{H}$  также задаются  $S$ -измеримыми функциями, и также две функции, отличающиеся значениями на множестве меры нуль, эквивалентны. Классы эквивалентных функций для краткости будем называть *ковекторами*.

<sup>4)</sup> Конусы  $C(\Omega, S, I)$  сохраняют следующее фундаментальное свойство конусов  $C(\Omega)$ . Отношение двух мер (их производная Радона–Никодима) есть положительная конечная  $\sigma$ -измеримая функция на  $\Omega$ , определенная с точностью до значений на нуль-множестве. Эти функции образуют положительный конус соответствующего полуполя; см. [9]. По отношению к умножению на такую функцию конус  $C(\Omega, S, I)$  является однородным пространством.

<sup>\*)</sup> Теория вероятн. и ее примен. 1966. Т. 11, вып. 3. С. 483–494.

В [1] на многообразиях  $\mathbf{H}(\Omega, S, I)$  была введена естественная линейная связность с нулевыми кручением и кривизной. В этой геометрии в  $\mathbf{H}(\Omega, S, I)$  существуют геодезические линии, а также вполне геодезические подмногообразия. вполне геодезические поверхности  $\gamma$  (конечного числа измерений) оказываются специальными семействами экспоненциальных распределений и в канонической аффинной параметризации  $\mathbf{s} = (s^1, \dots, s^n)$  описываются семействами плотностей

$$\frac{dP_{\mathbf{s}}}{d\mu}(\omega) = p(\omega; \mathbf{s}) = p_0(\omega) \exp[\mathbf{s} \cdot \mathbf{q} - \Psi(\mathbf{s})], \quad (1)$$

где  $\mathbf{q} = (q_1(\omega), \dots, q_n(\omega))$  — направляющие ковекторы, образующие необходимую и достаточную статистику семейства,  $\mathbf{s} \cdot \mathbf{q} = s^i q_i$ ,  $\exp[\Psi(\mathbf{s})]$  — нормирующий делитель; см. [1]. Всюду далее, если только не оговорено противное, мы считаем, что в представлении (1) функции  $q_0(\omega) \equiv 1, q_1(\omega), \dots, q_n(\omega)$  линейно независимы в области  $\Omega$ , так что соответствие между набором  $\mathbf{s}$  параметров и плотностью  $p(\omega; \mathbf{s})$  взаимно однозначное, и  $\dim \gamma = n$ .

Семейство (1) описывает аналитическую поверхность  $\gamma$  в  $\mathbf{H}$  и однозначно определяется любыми своими  $1 + n$  точками общего положения.

ЛЕММА 1. *Всякие две канонические параметризации  $\mathbf{s}$  и  $\bar{\mathbf{s}}$  семейства  $\gamma$  связаны соотношениями*

$$\bar{\mathbf{q}} = A\mathbf{q} + \mathbf{b}, \quad \bar{\mathbf{s}} = A^* \mathbf{s} + \mathbf{r} \quad (2)$$

с постоянной матрицей  $A$ ,  $\det A \neq 0$ ,  $A^* = (A^T)^{-1}$ , и постоянными векторами  $\mathbf{b}$  и  $\mathbf{r}$ .

Класс вполне геодезических семейств  $\gamma$  и более общий класс всех экспоненциальных семейств

$$p(\omega; \theta) = p_0(\omega) \exp[\mathbf{s}(\theta) \cdot \mathbf{q}(\omega) - \Psi(\mathbf{s}(\theta))] \quad (3)$$

произвольных “подповерхностей” во вполне геодезических поверхностях  $\gamma$  представляют большой интерес, так как семейства (3) и (в классе семейств взаимно абсолютно непрерывных распределений) только они обладают конечным числом достаточных статистик; см. [2, 3].

**2.** Систематическая теория экспоненциальных семейств в канонических координатах построена в [4]<sup>1)</sup>. В настоящей работе на вполне геодезических поверхностях исследуется другая замечательная координатизация, в некотором смысле двойственная канонической<sup>2)</sup>. Мы будем называть ее натуральной, так как она имеет непосредственный статистический смысл, и систематически изложим ее свойства. Теорию, изложенную в [4], а также в [1], будем предполагать известной.

ОПРЕДЕЛЕНИЕ 1. Назовем параметризацию  $\mathbf{t}(P_{\mathbf{s}}) = \{t_j\}_{j=1}^n$  вполне геодезического семейства (1) *натуральной*, если

$$t_j = \mathbf{M}_{\mathbf{s}} q_j(\omega) = \int q_j(\omega) p(\omega; \mathbf{s}) \mu\{d\omega\}. \quad (4)$$

<sup>1)</sup>По моей вине в заметке [1] нет ссылки на монографию [4].

<sup>2)</sup>В частном случае экспоненциальных распределений на прямой, зависящих от одного числового параметра, такая перепараметризация применена еще в [5]; см. также [6].

Из формул (2) видно, что натуральная параметризация  $\gamma$  также определена с точностью ровно до произвольного невырожденного аффинного преобразования пространства параметров  $\mathbf{t} \rightarrow A\mathbf{t} + \mathbf{b}$ .

Для упрощения выкладок мы закрепим систему канонических координат, положив

$$p(\omega; \mathbf{s}) = p(\omega; 0) \exp[\mathbf{s} \cdot \mathbf{q}(\omega; 0) - \mathcal{E}(\mathbf{s})], \quad (5)$$

где  $P_0\{\cdot\}$  — любая внутренняя (см. [1, 4]) точка поверхности  $\gamma$ , которая принята за начало отсчета  $\mathbf{s} = 0$ ;

$$p(\omega; 0) = \frac{dP_0}{d\mu}(\omega); \quad q_j(\omega; \mathbf{s}) = \frac{\partial \ln p(\omega; \mathbf{s})}{\partial s^j}$$

— касательные к  $\gamma$  в точке  $P_{\mathbf{s}}$  ковекторы. Отвечающую представлению  $\gamma$  в форме (5) натуральную параметризацию  $\mathbf{t}$  будем называть *двойственной* к  $\mathbf{s}$ . Очевидно, начало отсчета этих координат  $\mathbf{t}$  также совпадает с точкой  $P_0$ .

Формулу (5) удобно прологарифмировать:

$$s^j q_j(\omega; 0) = \ln \frac{p(\omega; \mathbf{s})}{p(\omega; \mathbf{0})} + \mathcal{E}(\mathbf{s}). \quad (6)$$

Для любого касательного  $q_k(\omega; \mathbf{s})$  имеем  $\mathbf{M}_{\mathbf{s}} q(\omega; \mathbf{s}) = 0$ . Отсюда получается второе выражение для  $\mathcal{E}(\mathbf{s})$ :

$$\int \exp[s^j q_j(\omega; 0)] P_0\{d\omega\} = \mathcal{E}(\mathbf{s}) = - \int \ln \left[ \frac{p(\omega; \mathbf{s})}{p(\omega; \mathbf{0})} \right] P_0\{d\omega\}. \quad (7)$$

Если обозначить количество информации по Шеннону, Калбеку и Лейблеру (см. [7]) об отличии  $P'$  от  $P''$  через

$$I[P''|P'] = - \int \ln \left[ \frac{dP''}{dP'} \right] P'\{d\omega\}, \quad (8)$$

то вторую половину формулы (7) можно записать как

$$\mathcal{E}(\mathbf{s}) = I[P_{\mathbf{s}}|P_0].$$

**ЛЕММА 2.** Если натуральная параметризация  $\mathbf{t}$  двойственна к канонической параметризации  $\mathbf{s}$  и  $\mathbf{q}(\omega) = \mathbf{q}(\omega, 0) + \mathbf{b}$ , то

$$t_k(\mathbf{s}) = \mathbf{M}_{\mathbf{s}} q_k(\omega; 0) = \frac{\partial}{\partial s^k} \mathcal{E}(\mathbf{s}), \quad \frac{\partial t_k}{\partial s^j} = \frac{\partial^2 \mathcal{E}}{\partial s^j \partial s^k} = \text{Cov}_s[q_k(\omega), q_j(\omega)]. \quad (9)$$

В [4] (см. также [1]) доказано, что значения  $n$ -мерного аффинного канонического параметра  $\mathbf{s}$ , которым соответствуют плотности семейства  $\gamma$ , заполняют некоторое выпуклое множество  $G^\gamma$  аффинного пространства параметров.

**ЛЕММА 3.** Внутри  $G^\gamma$  соответствие (4) между канонической и натуральной параметризациями является в обе стороны однозначным<sup>3)</sup> и аналитическим.

<sup>3)</sup> Поэтому для краткости будем писать  $p(\omega; \mathbf{t})$  и  $P_{\mathbf{t}}\{d\omega\}$  вместо  $p(\omega; \mathbf{s}(\mathbf{t}))$  и  $P_{\mathbf{s}(\mathbf{t})}\{d\omega\}$ .

ТЕОРЕМА 1. Если натуральная параметризация  $\mathbf{t}$  вполне геодезического семейства  $\gamma$  двойственна к канонической параметризации  $\mathbf{s}$ , то переход от переменных  $s$  и функции  $I[P_{\mathbf{s}}|P_0]$  к переменным  $t$  и функции  $I[P_0|P_{\mathbf{s}}]$  является преобразованием Лежандра

$$\begin{aligned} s^k t_k &= I[P_{\mathbf{s}}|P_0] + I[P_0|P_{\mathbf{s}}], \\ t_k &= \frac{\partial}{\partial s^k} I[P_{\mathbf{s}}|P_0], \quad s^k = \frac{\partial}{\partial t_k} I[P_0|P_{\mathbf{s}}]. \end{aligned} \quad (10)$$

ЛЕММА 4. Для пары двойственных параметризаций  $\mathbf{s} = \mathbf{s}(\mathbf{t})$  и  $\mathbf{t} = \mathbf{t}(\mathbf{s})$  реперы  $\left\{ \frac{\partial \ln p(\omega; \mathbf{t})}{\partial t_j} \right\}_{j=1}^n$  и  $\left\{ \frac{\partial \ln p(\omega; \mathbf{s})}{\partial s^k} \right\}_{k=1}^n$  в касательном в точке  $P$  копространстве измеримых функций  $f$ ,  $\mathbf{M}f = \int f(\omega) P\{d\omega\} = 0$  (см. [1]), образуют биортogonalную систему

$$\mathbf{M}_P \frac{\partial \ln p}{\partial t_j} \frac{\partial \ln p}{\partial s^k} = \int \frac{\partial \ln p(\omega; \mathbf{s})}{\partial s^k} \frac{\partial}{\partial t_j} \mathbf{P}_{\mathbf{t}}\{d\omega\} = \delta_k^j \quad (11)$$

относительно евклидовой метрики  $(f, f)_P = \mathbf{M}_P f^2 = \mathbf{D}_P f$ .

Элементы якобиана перехода от  $\mathbf{s}$  к  $\mathbf{t}$  суть

$$\frac{\partial t_j}{\partial s^k} = \mathbf{M}_{\mathbf{s}} \frac{\partial \ln p(\omega; \mathbf{s})}{\partial s^j} \frac{\partial \ln p(\omega; \mathbf{s})}{\partial s^k} = v_{jk}(\mathbf{s}), \quad (12)$$

а якобиана обратного перехода от  $\mathbf{t}$  к  $\mathbf{s}$  —

$$\frac{\partial s^k}{\partial t_j} = \mathbf{M}_{\mathbf{s}} \frac{\partial \ln p(\omega; \mathbf{t})}{\partial t_j} \frac{\partial \ln p(\omega; \mathbf{t})}{\partial t_k} = w^{jk}(\mathbf{t}). \quad (13)$$

**3.** Необходимая и достаточная векторная статистика  $\tau(\omega) = \mathbf{q}(\omega; 0)$  задает в силу (4) несмещенную оценку параметра  $\mathbf{t}$ . Дифференцируя (6):

$$\frac{\partial}{\partial s^k} \ln \frac{p(\omega; \mathbf{s})}{p(\omega; 0)} = q_k(\omega; 0) - t_k(\mathbf{s}), \quad (14)$$

убеждаемся, что решением системы уравнений наибольшего правдоподобия будет оценка  $\mathbf{t}^* = \tau(\omega)$ . Однако примеры показывают, что это решение уравнения правдоподобия может быть формальным.

ПРИМЕР 1. Семейство распределений с плотностью

$$p(x; \mathbf{s}) = \frac{1}{B_m(\mathbf{s})} \frac{1}{x^m} \exp[\mathbf{s}x], \quad (15)$$

заданное на луче  $x \geq 1$  и зависящее от индекса  $m > 2$ . Аффинный параметр этого семейства изменяется в пределах  $-\infty < \mathbf{s} \leq 0$ , соответствующий<sup>4)</sup> натуральный — в пределах  $1 < \mathbf{t} \leq 1 + 1/(m-2)$ , а статистика  $\tau(x)$  — в пределах

<sup>4)</sup>Для упрощения формулы (15) она написана не в специальной форме (5), а в форме (1).

$1 < x = \tau(x) < \infty$ . При  $x > 1 + 1/(m - 2)$  статистика  $\tau(x)$  дает формальную оценку  $\mathbf{t}$ .

**ТЕОРЕМА 2.** *Параметризация  $\mathbf{t}$  гладкого семейства  $\gamma$  взаимно абсолютно непрерывных законов распределения  $P\{\cdot\}$  допускает совместно эффективную оценку  $\tau(\omega)$  тогда и только тогда, когда семейство является вполне геодезическим, а  $\mathbf{t}$  — его натуральная параметризация.*

**ЗАМЕЧАНИЕ.** Оценка  $\tau(\omega)$  при этом будет единственным решением уравнения наибольшего правдоподобия; она может быть формальной с положительной вероятностью.

**Доказательство теоремы 2.** Как известно (см. [8]), критерием эффективности оценки  $\tau$  является выполнение при каждом  $\mathbf{t}$  равенств

$$\frac{\partial}{\partial t_j} \ln p(\omega; \mathbf{t}) = w^{jk} [\tau_k(\omega) - t_k], \tag{16}$$

где  $w^{jk}(\mathbf{t})$  определены математическими ожиданиями (13). Статистики  $\tau_0(\omega) \equiv 1, \tau_1(\omega), \dots, \tau_n(\omega)$  будем считать линейно независимыми (в противном случае семейство  $\gamma$  координатизируется  $\mathbf{t} \rightarrow p(\omega; \mathbf{t})$  уже частью параметров  $t_j = M_{\mathbf{t}} \tau_j(\omega)$ , а остальные параметры и равенства из (16) для них можно отбросить). Интегрируя соотношения (16) вдоль любого кусочно гладкого пути, соединяющего  $\mathbf{t}$  с  $\mathbf{t}_0$ , получаем

$$\ln p(\omega; \mathbf{t}) = \ln p(\omega; \mathbf{t}_0) + s^k(\mathbf{t}) \tau_k(\omega) = \Phi(\mathbf{t}), \tag{17}$$

где

$$s^k(\mathbf{t}) = \int_{\mathbf{t}_0}^{\mathbf{t}} w^{jk} dt_j, \quad \Phi(\mathbf{t}) = \int_{\mathbf{t}_0}^{\mathbf{t}} w^{jk} t_k dt_j.$$

В силу линейной независимости функций  $\tau_k(\omega)$  представление (17) для  $\ln p(\omega; \mathbf{t})$  единственно. Поэтому отображение  $\mathbf{t} \rightarrow \mathbf{s}(\mathbf{t})$ , вытекающее из (17), однозначно. Обратное, набор  $s^k$  в (17) вполне определяет  $\ln p(\omega; \mathbf{t})$ , ибо если только плотность  $p(\omega, \mathbf{t}_0) \exp[s^k \tau_k(\omega)]$  вообще интегрируема, то она нормируется единственным образом. Поэтому отображение  $\mathbf{s} \rightarrow p(\omega; \mathbf{s}) \rightarrow \mathbf{t}$  не более чем однозначно, и соответствие  $\mathbf{t}$  и  $\mathbf{s}$  взаимно однозначно. В новых координатах  $s^k(\mathbf{t})$  формула (17) совпадает с (1). Теорема доказана.

**ПРИМЕР 2.** Семейство нормальных распределений на прямой  $-\infty < x < \infty$  с переменными средним  $a = s^{(1)}D$  и дисперсией  $D = -1/s^{(2)}$  задается семейством плотностей

$$\frac{1}{\sqrt{2\pi D}} \exp \left\{ -\frac{(x-a)^2}{2D} \right\} = \exp \{ s^{(1)}x + s^{(2)}x^2 - \Psi(\mathbf{s}) \}, \tag{18}$$

где  $\Psi(\mathbf{s}) = (\ln D)/2 - a^2/(2D)$ . Параметр  $s^{(1)}$  изменяется в пределах  $-\infty < s^{(1)} < \infty$ , параметр  $s^{(2)}$  — в пределах  $-\infty < s^{(2)} < 0$ . Минимальной статистикой семейства является  $\tau(x) = (x, x^2)$ , соответствующие натуральные параметры —  $t_1 = M_{\mathbf{s}} \xi = a$  и  $t_2 = M_{\mathbf{s}} \xi^2 = a^2 + D$ . Так как  $a$  и  $D$  выражаются через  $t_1$  и  $t_2$  квадратично, то совместно они эффективно не оцениваются<sup>5)</sup>. Однако при

<sup>5)</sup>Как известно, граница эффективности совместной оценки  $a$  и  $D$  определяется уточнением неравенства информации; см. [8].

каждом фиксированном  $a$  дисперсия  $D$  является натуральным параметром соответствующего однопараметрического геодезического семейства нормальных законов.

4. Ковекторы  $q_1(\omega), \dots, q_n(\omega)$ , образующие в (1) определяющую статистику семейства  $\gamma$ , задают измеримое отображение измеримого пространства  $(\Omega, S)$  в  $n$ -мерное аффинное пространство  $U_n$ . Очевидно (см. [4]), это отображение  $\mathbf{u} = \mathbf{q}(\omega)$  индуцирует в  $U_n$  геодезическое же семейство  $\hat{\gamma}$  плотностей распределений вероятностей

$$\hat{p}(\mathbf{u}; \mathbf{s}) = \hat{p}_0(u) \exp[\mathbf{s} \cdot \mathbf{u} - \Psi(\mathbf{s})] \quad (19)$$

относительно меры  $\hat{\mu}\{\cdot\}$ , где  $\hat{\mu}\{A\} = \mu\{\omega : \mathbf{q}(\omega) \in A\}$ . Это семейство  $\hat{\gamma}$ , определенное с точностью до аффинного преобразования пространства  $U_n$ , мы назовем *фазовым изображением* семейства  $\gamma$  или его, семейства  $\gamma$ , приведенной (редуцированной) формой. Мера  $\hat{\mu}$  и соответственно распределения  $\hat{P}_s\{du\}$  на  $U_n$  могут быть дискретными (например, в случае, когда  $\Omega$  состоит из конечного числа исходов), абсолютно непрерывными (по лебеговой мере), как в примере 1, или сингулярными, как в примере 2, когда вся мера сосредоточена на параболы  $(u_1)^2 = u_2$ . Наименьшее замкнутое выпуклое множество полной  $\hat{\mu}$ , а следовательно, и  $\hat{P}_s$ -меры, мы назовем *носителем статистики*  $\mathbf{q}(\omega)$  и обозначим  $\hat{C}_\gamma$ . В приведенных выше примерах  $\hat{C}_\gamma$  является многогранником для конечных  $\Omega$ , лучом  $u \geq 0$  для семейства примера 1 и внутренностью параболы  $u_2 \geq (u_1)^2$  для семейства нормальных законов.

5. Прежде чем перейти к свойствам условных распределений, возникающих в геодезических семействах, выделим класс многообразий  $\mathbf{H} = \mathbf{H}(\Omega, S, I)$ , на которых существование этих распределений гарантировано.

ОПРЕДЕЛЕНИЕ 2. Назовем распределение  $P \in \mathbf{H}$  *распределением типа Лебега–Розлина* или, для краткости, *конструктивным*, если существует измеримое по Лебегу отображение единичного отрезка  $E[0 \leq x \leq 1]$  в  $\Omega$ ,  $\omega = \varphi(x)$ , индуцирующее на  $S$  распределение  $P$ :

$$P\{A\} = \lambda\{x : \varphi(x) \in A\},$$

где  $\lambda\{\cdot\}$  — мера Лебега,  $\lambda\{B\} = \int_B dx = \int_0^1 \chi_B(x) dx$ .

Нетрудно показать, что если распределение  $P$  является конструктивным,  $P \in \mathbf{H}(\Omega, S, I)$ , то и любое другое  $Q \in \mathbf{H}$  и даже любое распределение  $R$  на  $(\Omega, S)$ , доминируемое  $P$ , также являются конструктивными. Поэтому можно говорить о *конструктивных многообразиях*  $\mathbf{H}(\Omega, S, I)$ , о *конструктивных совокупностях*  $\text{Card}(\Omega, S, I)$  распределений вероятностей, обращающихся в нуль на множествах из  $I$ , и просто о *конструктивных тройках*  $(\Omega, S, I)$ .

В [9] и [10] автором было показано, что совокупность  $\text{Car}(\Omega, S)$  всех распределений вероятностей на измеримом пространстве  $(\Omega, S)$  есть объект категории  $\text{CAR}$  всех таких совокупностей с системой марковских гомоморфизмов  $\Pi: \text{Car}(\Omega', S') \rightarrow \text{Car}(\Omega'', S'')$ ,

$$P^{(\Pi)}\{\cdot\} = \int_{\Omega'} \Pi\{\omega'; \cdot\} P\{d\omega'\}, \quad (20)$$

где  $\Pi\{\omega; A\}$  — переходная марковская мера.

**ЛЕММА 5.** *Класс всех конструктивных многообразий  $\mathbf{H}(\Omega, S, I)$  образует категорию CAPH с системой конструктивных марковских морфизмов.*

Аналогично образуют категорию CAPD с той же системой конструктивных марковских морфизмов “замыкания” многообразий  $\mathbf{H}(\Omega, S, I)$  — конструктивные совокупности  $\text{Card}(\Omega, S, I)$ . Для доказательства леммы заметим, что если  $P \in \mathbf{H}(\Omega', S', I')$  доминирует  $Q$  на  $(\Omega', S')$ , то и распределение  $P^{(\Pi)}\{\cdot\}$  на  $(\Omega'', S'')$ , построенное по (20), доминирует  $Q^{(\Pi)}\{\cdot\}$ , образованное также по правилу (20). Таким образом, марковский гомоморфизм переводит каждое полное многообразие  $\mathbf{H}(\Omega', S', I')$  в полное же многообразие  $\mathbf{H}(\Omega'', S'', I'')$  с некоторым идеалом  $I''$ . (Хотя, может статься, ни при одном  $\omega'$  распределение  $\Pi\{\omega'; \cdot\}$  не принадлежит  $\mathbf{H}(\Omega'', S'', I'')$ .) Далее, когда распределение  $P\{d\omega'\}$  конструктивно и переходная мера  $\Pi\{\omega', A\}$  при каждом фиксированном  $\omega'$  есть конструктивное распределение на  $(\Omega'', S'')$ , то нетрудно построить конструкцию меры  $P^{(\Pi)}$ .

Замечательно, что последнее утверждение можно обратить. По известной теореме Колмогорова (см. [11]) для конструктивного распределения  $P\{\cdot\} \in \mathbf{H}(\Omega, S, I)$  всегда существует *конструктивный вариант условного распределения*  $P\{d\omega | \xi(\omega) = x\}$  на  $(\Omega, S)$  относительно отображения  $\xi(\omega)$  в единичный отрезок  $E[0 \leq x \leq 1]$ , а стало быть, и относительно значений любой случайной величины или нескольких случайных величин. (Заметим, что  $I$  не обязано быть идеалом нуль-множеств для условных распределений.)

Конструктивное задание какого-либо распределения довольно часто применяется как в теории<sup>6)</sup> (см. [13, 14]), так и при моделировании случайных величин (см. [15]). Поэтому выделение категории конструктивных совокупностей распределений и возникающего отношения конструктивной статистической эквивалентности представляется целесообразным. Впрочем, как следует из [16], некоторые из наших последующих предложений справедливы и в более широком классе совершенных распределений.

**6. ЛЕММА 6.** *Если система условных распределений  $P\{\cdot | \mathbf{q}(\omega) = \mathbf{u}\}$  существует на  $S$  хотя бы для одного  $P \in \gamma$ , то она существует при всех  $s \in G^\gamma$  и не зависит от  $s$ :*

$$P_s\{\cdot | \mathbf{q}(\omega) = \mathbf{u}\} = \Pi\{\mathbf{u}; \cdot\}.$$

**ЗАМЕЧАНИЕ.** Условие леммы заведомо выполнено, когда семейство  $\gamma$  принадлежит конструктивному многообразию  $\mathbf{H}(\Omega, S, I)$ .

Отсюда и из формулы (19) вытекает, что конструктивное геодезическое семейство  $\gamma$  распределений вероятностей можно представить в виде

$$P_s\{A\} = \int_{U_n} \hat{p}(\mathbf{u}; s) \Pi\{\mathbf{u}; A\} \hat{\mu}\{d\mathbf{u}\}. \quad (21)$$

Эта формула показывает, что конструктивное геодезическое семейство  $P_s\{d\omega\}$  статистически (конструктивно) эквивалентно своему фазовому изображению  $P_s\{d\mathbf{u}\}$ . Таким образом, два геодезических семейства статистически эквивалентны тогда и только тогда, когда статистически эквивалентны их редуцированные

<sup>6)</sup>Меры типа Лебега–Рохлина были систематически изучены в фундаментальной работе [12] в первую очередь в связи с приложениями к теории динамических систем.

формы. В книге [4] (см. также [1]) доказано, что нормирующий делитель  $a(\mathbf{z}) = \exp \Psi(\mathbf{z})$  — аналитическая функция внутри цилиндрической области  $\text{Re } \mathbf{z} \in G^\gamma$ ,  $\text{Re } \mathbf{z} = (\text{Re } z^1, \dots, \text{Re } z^n)$ . Легко проверить, что функция

$$\varphi_{\mathbf{s}}(\sigma) = \frac{a(\mathbf{s} + i\sigma)}{a(\mathbf{s})} = \exp [\Psi(\mathbf{s} + i\sigma) - \Psi(\mathbf{s})]$$

есть характеристическая функция случайного вектора  $q(\omega)$  с распределением случая  $P_{\mathbf{s}}\{d\omega\}$  (см. [17]), т.е. преобразование Фурье меры  $\widehat{P}_{\mathbf{s}}\{d\omega\}$  на  $U_n$ .

**ТЕОРЕМА 3.** *Два вполне геодезических семейства  $\gamma_1$  и  $\gamma_2$  конструктивных распределений вероятностей статистически эквивалентны тогда и только тогда, когда при соответствующей одинаковой канонической параметризации  $\mathbf{s}$  в некоторой открытой области  $\Psi_1(\mathbf{s}) = \Psi_2(\mathbf{s})$ .*

При этом эквивалентность конструктивна и  $\Psi_1(\mathbf{s}) = \Psi_2(\mathbf{s})$  всюду на  $G^\gamma$ .

Эта теорема является бесконечномерным аналогом, правда, только для геодезических семейств теоремы 3 работы [9] о строении эквивалентных семейств распределений на пространствах  $\Omega$  с конечным числом исходов.

**ТЕОРЕМА 4.** *Если у конструктивного геодезического семейства  $\gamma$ ,  $\dim \gamma = n$ , каноническая и натуральная параметризация совпадают, то семейство  $\gamma$  конструктивно эквивалентно семейству  $\widehat{\gamma}$   $n$ -мерных нормальных законов с постоянной матрицей дисперсий, и параметром — вектором средних значений. Обратное утверждение также справедливо.*

Для этих семейств нормальных законов матрица  $\mathbf{D}$  дисперсий определяет единственную естественную риманову метрику, превращающую пространство параметров в евклидово,  $\|\mathbf{u}\|^2 = (\mathbf{D}\mathbf{u}, \mathbf{u})$ . Такие семейства оказываются единственными поверхностями в  $\mathbf{H}$  постоянной нулевой внутренней кривизны, однородно вложенными в конус  $\mathbf{C}(\Omega, S, I)$  (см. [1]), так сказать, имеющими постоянную внешнюю кривизну. Заметим, что на множествах  $\Omega$  с конечным или счетным числом всех исходов таких поверхностей не существует.

**7.** Пусть  $\Omega_1 \subset \Omega$ ,  $\Omega_1 \in S$  и  $\Omega \notin I$ . Скажем, что  $\overline{\mu}\{\cdot\} = \text{Pr } \mu\{\cdot\}$  — проекция на  $\Omega_1$  меры  $\mu$  на  $\Omega$ , если  $\overline{\mu}$  определено на  $S$  равенством  $\overline{\mu}\{A\} = \mu\{A \cap \Omega_1\}$ . Далее,  $\Omega_1$ -усечением произвольного распределения  $P\{\cdot\}$  на  $(\Omega, S)$  будем называть распределение  $P^{(1)}\{\cdot\}$  на  $(\Omega, S)$ , задающееся условными вероятностями

$$P^{(1)}\{C\} = P\{C|\Omega_1\} = \frac{P\{C \cap \Omega_1\}}{P\{\Omega_1\}}.$$

При  $\Omega_1 \notin I$  операция  $\Omega_1$ -усечения задает нелинейный гомоморфизм многообразия  $\mathbf{H}(\Omega, S, I)$  в  $\mathbf{H}(\Omega, S, I_1)$ , где новый идеал  $I_1$  нуль-множеств образован всеми множествами вида  $C_1 \cup C_2$ ,  $C_1 \in I$ ,  $C_2 \subset \Omega \setminus \Omega_1$ .

**ТЕОРЕМА 5.**  $\Omega_1$ -усечения распределений  $P_{\mathbf{s}}\{\cdot\}$  вполне геодезического семейства  $\gamma \subset \mathbf{H}(\Omega, S, I)$  с плотностью (1) задаются по мере  $\mu$  плотностями

$$p^{(1)}(\omega; \mathbf{s}) = p_0(\omega) \chi(\omega) \exp[\mathbf{s} \cdot \mathbf{q}(\omega) - \Psi(\mathbf{s}) - \ln P_{\mathbf{s}}\{\Omega_1\}] \quad (22)$$

и образуют геодезически выпуклую область вполне геодезического семейства  $\gamma_1 \subset \mathbf{H}(\Omega, S, I_1)$  с той же канонической параметризацией

$$p^{(1)}(\omega; \mathbf{s}) = p_0^{(1)}(\omega) \exp[\mathbf{s} \cdot \mathbf{q}(\omega) - \Psi(\mathbf{s})],$$

где  $p_0^{(1)}(\omega) = p_0(\omega)$ , если плотность рассматривается по мере  $\bar{\mu} = \text{Pr} \mu$  и  $p_0^{(1)}(\omega) = p_0(\omega)\chi(\omega)$  для плотности по мере  $\mu$ ,  $\chi(\omega)$  — индикатор множества  $\Omega_1$ . Соответствующие натуральные параметризации связаны соотношением

$$t_j^{(1)}(\mathbf{s}) = t_j(\mathbf{s}) + \frac{\partial}{\partial s^j} \ln P_{\mathbf{s}}\{\Omega_1\}. \quad (23)$$

ЗАМЕЧАНИЕ.  $\dim \gamma_1 \leq \dim \gamma$ , и лишь когда функции  $q_0 \equiv 1, q_1, \dots, q_n$  линейно независимы не только в  $\Omega$ , но и в  $\Omega_1$ , то  $\dim \gamma_1 = \dim \gamma$ . В последнем случае параметры  $\mathbf{s}$  и  $\mathbf{t}$  задают в  $\gamma_1$  систему координат. Эффективную оценку нового натурального параметра  $\mathbf{t}$  задает та же статистика  $\mathbf{q}(\omega)$  или  $I_1$ -эквивалентная ей статистика  $\mathbf{g}(\omega) = \mathbf{q}(\omega)\chi(\omega)$ . Теорема является незначительным уточнением результата Тьюки [18].

Формула (21) показывает, что операцию усечения можно истолковать как операцию “проектирования”  $\mathbf{H}(\Omega, S, I)$  на его “грань”  $\mathbf{H}(\Omega, S, I_1)$  пучком ориентированных параллельных геодезических по направлению ковектора  $\chi(\omega)$ . Конечно, можно также рассматривать  $P\{C|\Omega_1\}$  только для  $C \subseteq \Omega_1$  на алгебре  $S_1 = S \cap \Omega_1$  подмножеств пространства  $\Omega_1$  с идеалом нуль-множеств  $I^{(1)} = I \cap \Omega_1$ . Соответствующая перефразировка теоремы 5 справедлива.

8. Пусть теперь  $\Omega = \Omega_1 \times \Omega_2$  — множество пар  $(\omega', \omega'')$  исходов  $\omega' \in \Omega_1, \omega'' \in \Omega_2, S = S_1 \otimes S_2$  — произведение соответствующих  $\sigma$ -алгебр,  $I = I_1 \circ I_2$  — возникающий в  $S$  идеал нуль-множеств. Назовем *внешним произведением*  $\gamma = \gamma_1 \times \gamma_2$  семейства  $\gamma_1 \subset \mathbf{H}(\Omega_1, S_1, I_1)$  распределений  $P_{\theta_1}^{(1)}$  и семейства  $\gamma_2 \subset \mathbf{H}(\Omega_2, S_2, I_2)$  распределений  $P_{\theta_2}^{(2)}$  семейство распределений  $P_{\theta} = P_{\theta_1}^{(1)} \times P_{\theta_2}^{(2)}$  на  $(\Omega, S)$ .

ЛЕММА 7. *Внешнее произведение вполне геодезических семейств  $P_{\mathbf{s}_1}^{(1)}$  и  $P_{\mathbf{s}_2}^{(2)}$  с каноническими параметрами  $\mathbf{s}_1$  и  $\mathbf{s}_2$  есть вполне геодезическое семейство  $P_{\mathbf{s}} \subset \mathbf{H}(\Omega, S, I)$  с каноническим параметром  $\mathbf{s} = (\mathbf{s}_1, \mathbf{s}_2)$ , плотностью по мере  $\mu = \mu_1 \times \mu_2$*

$$p(\omega; \mathbf{s}) = \bar{p}(\omega', \mathbf{s}_1) \cdot \bar{\bar{p}}(\omega'', \mathbf{s}_2) = \bar{p}_0(\omega') \bar{\bar{p}}_0(\omega'') \exp [\mathbf{s}_1 \cdot \bar{\mathbf{q}}(\omega') + \mathbf{s}_2 \cdot \bar{\bar{\mathbf{q}}}(\omega'') - \Psi^{(1)}(\mathbf{s}_1) - \Psi^{(2)}(\mathbf{s}_2)] \quad (24)$$

и соответствующей натуральной параметризацией

$$\mathbf{t} = (\bar{\mathbf{t}}(\mathbf{s}_1), \bar{\bar{\mathbf{t}}}(\mathbf{s}_2)).$$

Когда два семейства распределений  $P_{\theta}^{(1)}$  на  $(\Omega_1, S_1, I_1)$  и  $P_{\theta}^{(2)}$  на  $(\Omega_2, S_2, I_2)$  координатизированы одним и тем же (векторным) параметром, можно рассмотреть их (внутреннее) произведение  $P_{\theta} = P_{\theta}^{(1)} \times P_{\theta}^{(2)}$  — распределение из  $\mathbf{H}(\Omega, S, I)$ . Параметр пробегает только те значения, при которых определено как  $P_{\theta}^{(1)}$ , так и  $P_{\theta}^{(2)}$ . Именно так описывается, например, схема последовательных независимых испытаний:  $\Omega^N = \times_{i=1}^N \Omega_i, \Omega_i = \Omega, \Omega^N \ni \varepsilon_N = (\omega^{(1)}, \dots, \omega^{(N)})$ .

ТЕОРЕМА 6. *Произведение вполне геодезических семейств  $\gamma_1 \subset \mathbf{H}(\Omega_1, S_1, I_1)$  и  $\gamma_2 \subset \mathbf{H}(\Omega_2, S_2, I_2)$  с одинаковыми каноническими параметризациями  $\mathbf{s}$  является*

ся вполне геодезическим семейством  $\gamma \subset \mathbf{H}(\Omega, S, I)$ , определенным при значениях канонического параметра  $\mathbf{s} \in G^\gamma = G^{\gamma_1} \cap G^{\gamma_2}$ . Плотности по мере  $\mu = \mu_1 \times \mu_2$  распределений семейства равняются

$$p(\omega; \mathbf{s}) = \bar{p}(\omega'; \mathbf{s}) \bar{\bar{p}}(\omega''; \mathbf{s}) = \bar{p}_0(\omega') \bar{\bar{p}}_0(\omega'') \exp [\mathbf{s} \cdot \{\bar{\mathbf{q}}(\omega') + \bar{\bar{\mathbf{q}}}(\omega'')\} - \{\Psi_1(\mathbf{s}) + \Psi_2(\mathbf{s})\}], \quad (25)$$

соответствующие натуральные параметры связаны соотношением

$$\mathbf{t}(\mathbf{s}) = \bar{\mathbf{t}}(\mathbf{s}) + \bar{\bar{\mathbf{t}}}(\mathbf{s}).$$

Рассмотрим подробнее схему последовательных независимых испытаний  $P_\theta^{(N)}\{d\varepsilon\} = \times_{i=1}^N P_\theta\{d\omega^{(i)}\}$ ,  $\varepsilon = (\omega^{(1)}, \dots, \omega^{(N)})$ . Если  $\mathbf{t}(\mathbf{s})$  — натуральная параметризация семейства  $P_{\mathbf{s}}\{d\omega\}$ , то

$$\mathbf{t}^{(N)}(\mathbf{s}) = N\mathbf{t}(\mathbf{s}),$$

где  $\mathbf{t}^{(N)}(\mathbf{s})$  — натуральная параметризация семейства  $P_{\mathbf{s}}^{(N)}\{d\varepsilon\}$ , соответствующая той же канонической параметризации  $\mathbf{s}$ . Статистика  $\mathbf{q}^{(N)}(\varepsilon) = \mathbf{q}(\omega^{(1)}) + \dots + \mathbf{q}(\omega^{(N)})$ , фазовая мера  $\hat{P}_{\mathbf{s}}^{(N)}\{du\}$  есть  $N$ -я композиция меры  $\hat{P}_{\mathbf{s}}\{du\}$ , носитель  $\hat{G}_\gamma^{(N)} = N\hat{G}_\gamma$ . Наоборот, если фиксировать натуральную параметризацию  $\mathbf{t}$ , то  $\mathbf{s}^{(N)}(\mathbf{t}) = N\mathbf{s}(\mathbf{t})$ . Поэтому для изучения асимптотики по  $N$  в точке  $P \in \gamma$  надо выбрать ее за начальную, взять две системы двойственных координат и увеличивать в  $\sqrt{N}$  раз масштаб как по  $\mathbf{s}$ , так и по  $\mathbf{t}$ . При этом двойственность сохраняется. В новых координатах  $\bar{\mathbf{t}}$  и  $\bar{\bar{\mathbf{t}}}$  семейство  $\hat{P}_{\bar{\mathbf{s}}(\bar{\mathbf{t}})}^{(N)}$  асимптотически будет семейством нормальных законов со средним  $\mathbf{t}$  и постоянной матрицей  $\mathbf{V}$  дисперсий,  $\mathbf{V} = \|v_{jk}(0)\|$  (см. (12)). При этом дисперсия самой статистики  $\bar{\mathbf{q}}(\varepsilon)$  сходится к дисперсии предельного закона,  $\mathbf{D}_{\bar{\mathbf{s}}(\bar{\mathbf{t}})}\bar{\mathbf{q}}(\varepsilon) = \mathbf{V} + O(N^{-1/2})$ .

**9.** Рассмотрим еще один пример.

**ПРИМЕР 3.** Двупараметрическое семейство усеченных нормальных законов с “параметрами”  $a$  и  $D = \sigma^2$ . Область усечения  $\Omega_1 \subset [-\infty < x < \infty]$ ,

$$\frac{1}{P_{a,D}\{\Omega_1\}\sqrt{2\pi D}} \exp \left\{ -\frac{(x-a)^2}{2D} \right\} = \exp [s^{(1)}x + s^{(2)}x^2 - \Psi(\mathbf{s})]. \quad (26)$$

По-прежнему,  $a = s^{(1)}D$ ,  $D = -1/s^{(2)}$ , — минимальная статистика семейства  $\tau = (x, x^2)$ , соответствующие натуральные параметры  $t_1 = \mathbf{M}_{\mathbf{s}}\xi$  и  $t_2 = \mathbf{M}_{\mathbf{s}}\xi^2$ , их совместно-эффективные оценки по независимым наблюдениям  $\xi^{(1)}, \dots, \xi^{(N)}$  суть  $N^{-1} \sum_i \xi^{(i)}$  и  $N^{-1} \sum_i [\xi^{(i)}]^2$ . Однако теперь связь натуральных параметров  $t_1$  и  $t_2$  с  $a$  и  $D$  становится гораздо сложнее,  $a$  и  $D$  теряют статистический смысл. В частности, когда область усечения конечна,  $G^\gamma = [-\infty < s^{(1)}, s^{(2)} < \infty]$ , становятся возможными отрицательные значения  $D$ . При  $\Omega_1 = [b \leq x \leq c]$  выписываются явные формулы (ср. [19])

$$t_1 = a - \frac{D}{J} [E(c) - E(b)], \quad t_2 = D + a^2 - \frac{D}{J} [(c+a)E(c) - (b+a)E(b)].$$

Здесь

$$E(z) = E(z; a, D) = \exp \left\{ -\frac{(z-a)^2}{2D} \right\}, \quad J = J(b, c; a, D) = \int_b^c E(z) dz.$$

При  $D > 0$ ,  $s^{(2)} < 0$  функция  $J$  выражается через интеграл вероятностей erf.

10. В заключение остановимся коротко на разумной интерпретации формальных значений оценки натурального параметра. Если обозначить  $G_\gamma$  область изменения параметра  $\mathbf{t} = \mathbf{t}(\mathbf{s})$  при  $\mathbf{s} \in G^\gamma$ , то решение уравнения наибольшего правдоподобия  $\mathbf{t}^* = \mathbf{q}(\omega^*)$  формально, когда  $\mathbf{t}^* \in \widehat{G}_\gamma$ , но  $\mathbf{t}^* \notin G_\gamma$ . Обычно в этом случае за  $\mathbf{s}^*$  берут “наиболее правдоподобное”  $\mathbf{s}$ , при котором достигается максимум плотности  $p(\omega^*, \mathbf{s})$ , или, что то же, максимум функции правдоподобия

$$\ln p(\omega^*; \mathbf{s}) = \ln p(\omega^*, 0) = s^k q_k(\omega^*) - \Psi(\mathbf{s}).$$

В силу выпуклости  $\Psi(\mathbf{s})$  таких максимумов не более одного. Однако иногда этот максимум реализуется при бесконечном значении  $\mathbf{s}$ . Подробный разбор возникающей ситуации, связанной со строением границы носителя  $\widehat{G}_\gamma$ , будет сделан в другом месте. Отметим лишь, что в схеме последовательных независимых испытаний, когда истинное значение  $\mathbf{s}$  лежит строго внутри  $G^\gamma$ , вероятность получить формальную оценку экспоненциально убывает с ростом числа наблюдений; ср. [20], также [21].

#### СПИСОК ЛИТЕРАТУРЫ

1. Ченцов Н.Н. Геометрия многообразия распределений вероятностей // Докл. АН СССР. 1964. Т. 158, № 3. С. 543–546.
2. Koopman B.O. On distributions admitting a sufficient statistic // Trans. Amer. Math. Soc. 1936. V. 39. P. 399–409.
3. Дынкин Е.Б. Необходимые и достаточные статистики // Успехи матем. наук. 1951. Т. 6, вып. 1. С. 66–90.
4. Леман Э. Проверка статистических гипотез. — М.: Наука, 1964.
5. Girshick M.A., Savage L.J. Bayes and minimax estimates for quadratic loss functions // Proc. Second Berkeley Sympos. Math. Statist. and Prob. 1951. P. 53–73.
6. Kullback S. Information Theory and Statistics. — N.Y., 1959.
7. Kullback S., Leibler R.A. On information and sufficiency // Ann. Math. Statist. 1951. V. 22, № 1. С. 79–86.
8. Большев Л.Н. Уточнение неравенства Крамера–Рао // Теория вероятн. и ее примен. 1961. Т. 6, вып. 3. С. 319–326.
9. Ченцов Н.Н. Категории математической статистики // Докл. АН СССР. 1965. Т. 164, № 3. С. 511–514.
10. Ченцов Н.Н. Категории математической статистики (резюме доклада на ММО) // Успехи матем. наук. 1965. Т. 20, вып. 4. С. 204–205.
11. Колмогоров А.Н. Основные понятия теории вероятностей. — М.–Л., 1936.
12. Рохлин В.А. Об основных понятиях теории меры // Матем. сб. 1949. Т. 25, № 1. С. 107–150.
13. Пэли Р., Винер Н. Преобразование Фурье в комплексной области. — М.: Наука, 1964.
14. Винер Н. Нелинейные задачи в теории случайных процессов. — М.: ИЛ, 1961.
15. Гельфанд И.М., Фролов А.С., Ченцов Н.Н. Вычисление континуальных интегралов методом Монте-Карло // Изв. Высш. учебн. заведений. Сер. матем. 1958. Т. 5, № 6. С. 32–45.

16. Сазонов В.В. О совершенных мерах // Изв. АН СССР. Сер. матем. 1962. Т. 26, № 3. С. 391–414.
17. Chernoff H. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations // Ann. Math. Statist. 1952. V. 23. P. 493–507.
18. Tukey J.W. Sufficiency, truncation and selection // Ann. Math. Statist. 1949. V. 20, № 3. P. 309–311.
19. Doss S.A.D.C. On uniqueness and maxima of the roots of likelihood equations under truncated and censored sampling from normal populations // Sankhya. 1962. A24, № 4. P. 355–362.
20. Wald A. Note on consistency of the maximum likelihood estimate // Ann. Math. Statist. 1949. V. 20. P. 595–601.
21. Ченцов Н.Н. Об асимптотической эффективности оценки наибольшего правдоподобия // Труды IV Всес. совещ. по теории вероят. и матем. статист. — Вильнюс, 1962. — С. 399–402.

## НЕСИММЕТРИЧНОЕ РАССТОЯНИЕ МЕЖДУ РАСПРЕДЕЛЕНИЯМИ ВЕРОЯТНОСТЕЙ, ЭНТРОПИЯ И ТЕОРЕМА ПИФАГОРА <sup>\*)</sup>

С. Кульбак и Р.А. Лейблер [1] ввели понятие информационного уклонения  $I[Q|P]$  распределения вероятностей  $Q\{d\omega\}$  от распределения вероятностей  $P\{d\omega\}$  на одном и том же измеримом пространстве  $(\Omega, S)$ :

$$I[Q|P] = - \int_{\Omega} \left[ \ln \frac{dQ}{dP}(\omega) \right] P\{d\omega\}, \quad (1)$$

обобщающее как понятие энтропии Шеннона, так и понятие энтропии в статистической физике. Обсуждение статистического смысла величины  $I[Q|P]$  дано в монографии [2]. Она всегда неотрицательна и, как правило, несимметрична. Как указано в [2] и [3], функция  $I[Q|P]$  играет фундаментальную роль в теории геодезических семейств — экспонентных семейств  $\gamma$  распределений вероятностей  $P_s\{d\omega\}$  с плотностью по фиксированной мере  $\mu$

$$\frac{dP_s}{d\mu}(\omega) = p(\omega; \mathbf{s}) = p_0(\omega) \exp[s^i q_j(\omega) - \Psi(\mathbf{s})], \quad (2)$$

$$\Psi(\mathbf{s}) = \ln \int_{\Omega} p_0(\omega) \exp[s^j q_j(\omega)] \mu\{d\omega\}. \quad (3)$$

Здесь  $\exp[\Psi(\mathbf{s})]$  — нормирующий делитель,  $\mathbf{s} = (s^1, \dots, s^n)$  — канонический аффинный параметр, статистики  $q_0(\omega) \equiv 1, q_1(\omega), \dots, q_n(\omega)$  предполагаются линейно независимыми.

<sup>\*)</sup> Матем. заметки. 1968. Т. 4, вып. 3. С. 323–332.

В категории совокупностей  $\text{Car}(\Omega, S)$  всех распределений вероятностей на измеримых пространствах  $(\Omega, S)$  с марковскими отображениями (описывающими вальдовские решающие правила) в качестве морфизмов (см. [4]) существует инвариантная аффинная связность<sup>1)</sup>. Она имеет нулевые кручение и кривизну. Ее вполне геодезическими поверхностями (в том числе линиями), т.е. аналогами  $n$ -мерных плоскостей (в том числе прямых), являются семейства (2). Их характеристическое свойство состоит в том, что логарифмы плотностей как функции  $\omega$  принадлежат  $(n + 1)$ -мерной плоскости  $\ln p_0(\omega) + s^j q_j(\omega) - \Psi(\mathbf{s}) \cdot 1(\omega)$ , “параллельной” направляющим статистикам  $q_1(\omega), \dots, q_n(\omega)$  и  $q_0(\omega) = 1(\omega) \equiv 1$ .

Каноническая аффинная параметризация семейства  $\gamma$  определена с точностью до произвольного невырожденного аффинного преобразования  $A$  координат  $s = As^*$  (см. [3]):

$$s^j = a_k^j s^{k*} + u^j, \quad q_j(\omega) = \alpha_j^k q_k^*(\omega) + b_j, \quad (4)$$

где  $a_k^j, \alpha_j^k, u^j, b_j$  не зависят от  $\omega$  и  $a_k^j \alpha_j^k = \delta_k^j$ . Если  $p_0(\omega)$  является плотностью распределения вероятностей, то

$$\Psi^*(\mathbf{s}^*) = \Psi(As^*) - a_k^j s^{k*} b_j - \Psi(u). \quad (5)$$

Введем еще натуральную параметризацию  $t$  геодезического семейства, отвечающую канонической

$$t_j(\mathbf{s}) = \mathbf{M}_s q_j(\omega) = \int_{\Omega} q_j(\omega) p(\omega; \mathbf{s}) \mu\{d\omega\} = \frac{\partial}{\partial s^j} \Psi(\mathbf{s}). \quad (6)$$

При преобразовании (4) системы канонических координат замена натуральных координат происходит по правилу

$$t_j = \alpha_j^k t_k^* + b_j,$$

которое получается, если в (4) перейти от статистик  $q_j(\omega)$  к их математическим ожиданиям.

Наиболее удобно представление (2), когда  $p_0(\omega) = p(\omega; 0)$  и за направляющие статистики выбраны

$$q_j(\omega) = \left. \frac{\partial \ln p(\omega; \mathbf{s})}{\partial s^j} \right|_{\mathbf{s}=0} = q_j(\omega; 0) \quad (7)$$

и, следовательно,  $t(0) = 0$ . Для приведения (2) к указанной стандартной форме достаточно сделать замену

$$q_j(\omega) = q_j^*(\omega) - \mathbf{M}_0 q_j(\omega), \quad u^j = 0, \quad a_k^j \alpha_j^k = \delta_k^j.$$

Нетрудно проверить, что при записи плотности (2) в стандартной форме

$$\Psi(s) = I[P_s | P_0]. \quad (8)$$

<sup>1)</sup>Эта связность была ранее введена нами в [9]. Она является инвариантной относительно категории марковских отображений (см. [4]) и притом единственной инвариантной аффинной связностью совокупностей  $\text{Car}(\Omega, S)$ .

Имеет место следующий замечательный факт (см. [3]). *Переход от переменных  $\mathbf{s} = (s^1, \dots, s^n)$  и выпуклой функции  $I[P_s|P_0]$  к переменным  $t = (t_1, \dots, t_n)$  и функции  $I[P_0|P_s(t)]$  является преобразованием Лежандра и<sup>2)</sup>*

$$s^j t_j(s) = I[P_s|P_0] + I[P_0|P_s]. \quad (9)$$

Распределения вероятностей  $P_{\mathbf{s}}$  семейства  $\gamma$  определены не при всех значениях аргумента  $\mathbf{s}$ , а только в некоторой выпуклой области  $G^\gamma$ , где интегралы (3) конечны (см. [6]). Чтобы не усложнять формулировок наших утверждений, будем предполагать, что семейство  $\gamma$  регулярно в следующем точном смысле.

Во-первых, выпуклая область  $G^\gamma$  имеет размерность  $n$ , и, во-вторых, каково бы ни было распределение вероятностей  $R\{\cdot\}$ , взаимно абсолютно непрерывное с любым  $P \in \gamma$ , в семействе  $\gamma$  существует распределение вероятностей  $P_\sigma\{\cdot\}$  с натуральными координатами

$$t_j(\sigma) = \int q_j(\omega) R\{d\omega\}. \quad (10)$$

Это предположение регулярности эквивалентно тому, что натуральный параметр  $\mathbf{t}$  пробегает всю выпуклую оболочку статистики  $\mathbf{q}(\omega) = (q_1(\omega), \dots, q_n(\omega))$  — наименьшее замкнутое выпуклое множество  $\hat{G}_\gamma$   $n$ -мерного пространства  $T$  такое, что  $P_0 \mathbf{q}^{-1}\{\hat{G}_\gamma\} = 1$ . В частности, когда семейство одномерно, условие регулярности означает, что

$$t(s) = \frac{\int q(\omega) p_0(\omega) \exp[sq(\omega)] \mu\{d\omega\}}{\int p_0(\omega) \exp[sq(\omega)] \mu\{d\omega\}}$$

принимает все значения в интервале

$$\text{vrai inf } q(\omega) < t < \text{vrai sup } q(\omega),$$

что, очевидно, выполняется, если обе эти грани конечны, но может быть неверно, если хоть одна из них бесконечна (см. [3, пример 1]).

**ЛЕММА 1.** *Пусть  $\gamma$  — геодезическое семейство распределений вероятностей,  $\mathbf{0}, \mathbf{s}, \mathbf{z} \in G^\gamma$  и натуральный параметр  $\mathbf{t}(\mathbf{z})$  конечен. Тогда*

$$I[P_{\mathbf{s}}|P_{\mathbf{z}}] = I[P_{\mathbf{s}}|P_{\mathbf{0}}] + I[P_{\mathbf{0}}|P_{\mathbf{z}}] - s^j t_j(\mathbf{z}). \quad (11)$$

**Доказательство.** Сделаем параллельный перенос системы канонических координат так, чтобы точка начала отсчета перешла из точки  $P_{\mathbf{0}}$  в точку  $P_{\mathbf{z}}$ . Тогда  $\mathbf{s} = \mathbf{s}^* + \mathbf{z}$ ,  $\mathbf{t} = \mathbf{t}^* + \mathbf{t}(\mathbf{z})$ ,  $\mathbf{q}(\omega; \mathbf{0}) = q^*(\omega) + \mathbf{t}(\mathbf{z})$ . Подставляя эти соотношения в (5), получаем  $\Psi^*(\mathbf{s}^*) = \Psi(\mathbf{s}) - (s^j + z^j)t_j(\mathbf{z}) - \Psi(\mathbf{z})$ .

Заменяя далее по (8) функцию  $\Psi^*(\mathbf{s}^*)$  на  $I[P_{\mathbf{s}}|P_{\mathbf{z}}]$ ,  $\Psi(\mathbf{z}) = I[P_{\mathbf{z}}|P_{\mathbf{0}}]$  и  $z^j t_j(\mathbf{z}) + I[P_{\mathbf{s}}|P_{\mathbf{0}}]$  на  $I[P_{\mathbf{0}}|P_{\mathbf{z}}]$  согласно (9), выводим (11).

<sup>2)</sup>Физическая энтропия, определяемая для гиббсовского канонического распределения  $P_{\mathbf{s}}$ , принадлежащего одномерному каноническому семейству (2), и есть  $I[P_{\mathbf{0}}|P_{\mathbf{s}}] = st(s) - I[P_{\mathbf{s}}|P_{\mathbf{0}}]$  (см. [5]).

СЛЕДСТВИЕ. Если  $\mathbf{0}, \sigma, \sigma + \mathbf{z} \in G^\gamma$  и  $\mathbf{t}(\sigma)$  конечно,

$$I[P_{\sigma+\mathbf{z}}|P_0] = I[P_{\sigma+\mathbf{z}}|P_\sigma] + I[P_\sigma|P_0] + z^j \frac{\partial}{\partial s^j} I[P_s|P_0]_{s=\sigma}.$$

На последнюю формулу можно смотреть как на своего рода “формулу сложения” специальной функции  $I[P_s|P_0]$ , ассоциированной с “поверхностью”  $\gamma$ . Здесь стоит отметить, что поле величин  $I[P_s|P_z]$  определяет геодезическое семейство  $\gamma$  с точностью до эквивалентности. Два конструктивных (см. [3]) геодезических семейства  $\gamma_1$  и  $\gamma_2$  распределений вероятностей соответственно  $P_\theta\{\cdot\}$  и  $Q_\theta\{\cdot\}$  в одинаковой (не обязательно канонической) гладкой параметризации эквивалентны тогда и только тогда, когда  $I[P_{\theta'}|P_\theta] = I[Q_{\theta'}|Q_\theta]$  для всех пар  $\theta', \theta$  в некоторой открытой области изменения аргумента.

Как мы уже отмечали в [7], всякая дважды дифференцируемая функция потерь  $L(Q; P)$ , инвариантная относительно категории статистических решений, является (с точностью до множителя) аналогом квадрата евклидова расстояния, вообще говоря, несимметричным. Для функции  $I[Q|P]$  эта аналогия заходит так далеко, что для нее справедлив несимметричный аналог равенства Пифагора в формулировке: “квадрат длины наклонной равен сумме квадрата длины перпендикуляра и квадрата длины проекции наклонной”.

ТЕОРЕМА 1. Пусть  $P\{\cdot\}$ ,  $Q\{\cdot\}$ ,  $R\{\cdot\}$  — взаимно абсолютно непрерывные распределения вероятностей на измеримом пространстве  $(\Omega, S)$ , и пусть

$$\int_{\Omega} \left[ \ln \frac{dP}{dQ}(\omega) \right] Q\{d\omega\} = \int_{\Omega} \left[ \ln \frac{dP}{dQ}(\omega) \right] R\{d\omega\}. \quad (12)$$

Тогда

$$I[P|R] = I[P|Q] + I[Q|R]. \quad (13)$$

Доказательство. Выпишем интеграл (1) для  $I[P|R]$  и преобразуем его, используя соотношение (12):

$$\begin{aligned} I[P|R] &= - \int_{\Omega} \left[ \ln \frac{dP}{dR}(\omega) \right] R\{d\omega\} = - \int_{\Omega} \left[ \ln \frac{dP}{dQ}(\omega) + \ln \frac{dQ}{dR}(\omega) \right] R\{d\omega\} = \\ &= - \int_{\Omega} \left[ \ln \frac{dP}{dQ}(\omega) \right] Q\{d\omega\} - \int_{\Omega} \left[ \ln \frac{dQ}{dR}(\omega) \right] R\{d\omega\} = I[P|Q] + I[Q|R]. \end{aligned}$$

Теорема доказана.

Будем рассматривать интеграл от  $S$ -измеримой функции  $f(\omega)$  по мере  $\nu\{d\omega\}$  на  $(\Omega, S)$  как внутреннее произведение функции  $f(\omega)$  (ковектора) и меры  $\nu\{d\omega\}$  (вектора)

$$\langle f, \nu \rangle = \int_{\Omega} f(\omega) \nu\{d\omega\}. \quad (14)$$

Тогда условие (12) есть условие “ортогональности” “направления”  $R - Q$  и “ко-направления”  $\ln(dP/dQ)$ . При такой интерпретации утверждение теоремы 1 есть (с точностью до постоянного множителя) точный несимметричный аналог теоремы Пифагора.

Удобно принять, что информационное уклонение  $I[P|Q]$  является аналогом половины квадрата расстояния, поскольку

$$I[Q|P] \approx \frac{1}{2} \int_{\Omega} \frac{[Q\{d\omega\} - P\{d\omega\}]^2}{P\{d\omega\}} = \frac{1}{2}(Q - P, Q - P)_P$$

с точностью до малых высшего порядка (см. [7]). Тогда формула (11) становится несимметричным аналогом теоремы косинусов, где

$$s^j t_j(\mathbf{z}) = \left\langle \ln \frac{dP_{\mathbf{s}}}{dP_0}, P_{\mathbf{z}} - P_0 \right\rangle \quad (15)$$

— внутреннее произведение ковектора  $\ln(dP_{\mathbf{s}}/dP_0)$ , отвечающего одной стороне треугольника  $P_0 P_{\mathbf{s}} P_{\mathbf{z}}$ , и вектора  $P_{\mathbf{z}} - P_0$ , отвечающего другой стороне.

Пусть  $\gamma$  — геодезическое семейство распределений вероятностей  $P_{\mathbf{s}}\{\cdot\}$  на  $(\Omega, S)$  и распределение  $R\{\cdot\}$  взаимно абсолютно непрерывно с распределениями  $P$ . Требуется отыскать “проекцию”  $R_{\gamma} = P_{\sigma}$  “точки”  $R$  на “плоскость”  $\gamma$ , минимизирующую отличие от  $R$ :

$$I[R_{\gamma}|R] = \inf_{P \in \gamma} I[P_{\mathbf{s}}|R]. \quad (16)$$

Назовем эту задачу “проектирования” *задачей* Б. В работах И.Н. Санова [8] и С. Кульбака [2] рассматривалась (при  $n = 1$ ) двойственная нашей задаче “проектирования”, которую мы назовем *задачей* А.

Пусть дано распределение  $P\{\cdot\}$  на  $(\Omega, S)$  и  $S$ -измеримые статистики  $q_1(\omega), \dots, q_n(\omega)$ , в непустом линейно выпуклом множестве — “гиперплоскости”  $\lambda$  распределений вероятностей  $R\{\cdot\}$ , взаимно абсолютно непрерывных с  $P$  и удовлетворяющих условиям

$$\int q_j(\omega) R\{d\omega\} = \tau_j, \quad j = 1, \dots, n. \quad (17)$$

Требуется найти распределение вероятностей  $P^{\lambda}$ , минимизирующее отличие  $P$ ,

$$I[P|P^{\lambda}] = \inf_{R \in \lambda} I[P|R]. \quad (18)$$

Покажем, что в регулярном случае решение обеих задач существует, единственно и задается распределением вероятностей  $P_{\sigma} \in \gamma$  с натуральными координатами

$$t_j(\sigma) = \tau_j = \int q_j(\omega) R\{d\omega\}. \quad (19)$$

**ТЕОРЕМА 2.** Пусть  $\gamma$  — регулярное геодезическое семейство (2) распределений вероятностей  $P_{\mathbf{s}}\{\cdot\}$  на  $(\Omega, S)$ , взаимно абсолютно непрерывных с распределением  $R\{\cdot\}$ .

Тогда решение задачи Б существует и единственно, а проекция  $R_{\gamma} = P_{\sigma}$  “точки”  $R$  на “плоскость”  $\gamma$  определяется своими натуральными координатами (19). При этом для любого  $P \in \gamma$

$$I[P|R] \geq I[P|R_{\gamma}], \quad (20)$$

где знак равенства может быть лишь при  $R_\gamma = R \in \gamma$ .

ТЕОРЕМА 3. Пусть  $P\{\cdot\}$  — распределение вероятностей на  $(\Omega, S)$ ,  $\lambda$  — непустая “гиперплоскость” (17) распределений вероятностей  $R\{\cdot\}$ , взаимно абсолютно непрерывных с  $P$ .

Тогда если регулярно геодезическое семейство  $\gamma$ , проходящее через  $P$  с направляющими статистиками  $q_1(\omega), \dots, q_n(\omega)$ , то решение задачи А существует и единственно, а проекция  $P^\lambda = P_\sigma$  “точки”  $P$  на “гиперплоскость”  $\lambda$  принадлежит  $\gamma$  и определяется своими натуральными координатами (19). При этом для любого  $R \in \lambda$

$$I[P|R] \geq I[P^\lambda|R], \tag{21}$$

где знак равенства может быть лишь при  $P^\lambda = P \in \lambda$ .

Доказательство теорем. Обозначим через  $Q$  распределение  $P_\sigma$  с натуральными координатами  $\tau_j$ , заданными (19). Оно существует в силу предположенной регулярности семейства  $\gamma$  и единственно в силу строгой выпуклости функции  $I[P_s|P_0]$ , гарантирующей взаимную однозначность  $s$  и  $t$  (см. [3]). Так как

$$\int q_j(\omega) P_\sigma\{d\omega\} = \tau_j = \int q_j(\omega) R\{d\omega\},$$

то условие (19) есть условие ортогональности ковектора  $\ln(dP/dQ)(\omega)$ , где  $P\{\cdot\} = P_s\{\cdot\} \in \gamma$ , и вектора  $R\{\cdot\} - Q\{\cdot\}$ . Так как информационное отличие всегда положительно и равно нулю лишь для совпадающих аргументов, то одновременно

$$\begin{aligned} I[P|R] &> I[Q|R] \quad \text{при} \quad P \neq P_\sigma = Q, \\ I[P|R] &> I[P|Q] \quad \text{при} \quad R \neq P_\sigma = Q, \end{aligned}$$

что дает оба утверждения теоремы 2 и оба утверждения теоремы 3.

Теорема 3 играет важную роль в теории больших отклонений случайных средних (см. [8]). Теорема 2 полезна в теории оптимальных оценок параметра закона распределения по наблюдениям. Рассмотрим еще несколько свойств “проектирования” на геодезическую поверхность, полезных для статистических приложений. Для такого проектирования справедливы аналоги известных из курса стереометрии теорем о двух и трех перпендикулярах.

ЛЕММА 2. Для того чтобы “точка”  $R_\gamma$  была проекцией  $R$  на  $n$ -мерную геодезическую поверхность  $\gamma$ , необходимо и достаточно, чтобы она была проекцией на  $n$  проходящих через  $R_\gamma$  геодезических линий общего положения.

ЛЕММА 3. Пусть  $\gamma$  есть  $k$ -мерная геодезическая поверхность, лежащая в  $n$ -мерной геодезической поверхности  $\Gamma$ . Тогда проекция  $R_\gamma$  “точки”  $R$  на  $\gamma$  является проекцией на  $\gamma$  проекции  $R_\Gamma$  на  $\Gamma$ .

Утверждения обеих лемм являются следствиями формулы (19) при специальном выборе направляющих статистик  $q_j(\omega)$ .

Когда семейство  $\gamma$  является нерегулярным или когда  $R$  и  $P$  не являются абсолютно непрерывными друг относительно друга, обе задачи проектирования сильно усложняются. В частном случае, когда мера  $R$  дискретна, задача Б приводится после “перенормировки”, сводящейся к вычитанию бесконечно большого постоянного слагаемого, к нахождению оценки наибольшего правдоподобия.

Сформулируем здесь только аналог предложения евклидовой геометрии о точке выпуклого множества, ближайшей к данной.

ЛЕММА 4. Пусть  $\gamma$  есть  $n$ -мерное геодезическое семейство распределений вероятностей  $P_s\{\cdot\}$ , и пусть  $C$  — произвольное ограниченное замкнутое выпуклое подмножество аффинного пространства параметров. Обозначим

$$F = \{P_s : s = C \cap G^\gamma\}.$$

Тогда для любого  $P$ -доминированного закона  $R\{\cdot\}$ , где

$$\inf_{P \in F} I[P_s|R] = I_F(R) < \infty,$$

существует и единственна “проекция”  $P_\sigma = P_F$ , минимизирующая  $I[P_s|R]$ ,

$$I[P_\sigma|R] = I_F(R).$$

Для всех законов  $P_\sigma \in F$  при  $R \notin F$

$$I[P_s|R] > I[P_\sigma|R].$$

Доказательство. Построим, согласно [9], “геодезическую оболочку”  $\Gamma$  точки  $R$  и “плоскости”  $\gamma$ . Семейство  $\Gamma$  будет  $(n+1)$ -мерным с направляющими статистиками  $q_1(\omega), \dots, q_n(\omega), q_{n+1}(\omega) = \ln(dR/dP_0)(\omega)$ . Функция  $\Psi_z(\mathbf{s}) = I[P_s|P_z]$  строго выпукла и непрерывна снизу (ср. [3]), а потому на компактном множестве  $C$  достигает своей нижней грани и притом лишь в одной точке  $\sigma$ ,  $P_\sigma = Q$ . Отсюда при  $P_z = R$  получаем первое утверждение леммы. Далее перенесем начало отсчета канонического параметра в произвольную точку  $P \in \gamma$ , где  $I[P|R] < \infty$ . Пусть  $R = P_z$ ,  $Q = P_\sigma$ . Соединим  $P$  с  $Q$  отрезком геодезической  $P_{\mathbf{s}(\alpha)} \subset \gamma$ ,  $\mathbf{s}(\alpha) = \alpha\sigma$ . По условию минимума  $(d/d\alpha)I[P_{\mathbf{s}(\alpha)}|R]_{\alpha=1} \leq 0$ . Дифференцируя по  $\alpha$  выражение для  $I[P_{\mathbf{s}(\alpha)}|P_z]$ , расписанное в форме (11), получаем  $\sigma^j t_j(\sigma) - \sigma^j t_j(z) \leq 0$ . Положим в (11)  $\mathbf{s} = \sigma$  и воспользуемся выведенным неравенством и тождеством (9):

$$I[P|Q] = \sigma^j t_j(\sigma) - I[P_\sigma|P_0] < \sigma^j t_j(\mathbf{z}) - I[P_\sigma|P_0] + I[P_\sigma|P_z] = I[P|R],$$

что и требовалось.

В заключение заметим, что для семейства  $\mathcal{N}$  многомерных нормальных законов с постоянной матрицей вторых центральных моментов и переменным средним наши предложения превращаются в теоремы евклидовой геометрии. Можно доказать, что это свойство является характеристическим свойством семейств  $\mathcal{N}$ .

#### СПИСОК ЛИТЕРАТУРЫ

1. Kullback S., Leibler R.A. On information and sufficiency // Ann. of Math. Stat. 1951. V. 22, № 1. P. 79–86.
2. Кузьбак С. Теория информации и статистика. — М., 1967.
3. Ченцов Н.Н. К систематической теории экспоненциальных семейств распределений // Теория вероятн. и ее примен. 1966. Т. 11, вып. 3. С. 483–494.

4. Ченцов Н.Н. Категории математической статистики // Докл. АН СССР. 1965. Т. 164, № 3. С. 511-514.
5. Хинчин А.Я. Математические основания статистической механики. — М.-Л., 1943.
6. Леман Э. Проверка статистических гипотез // М., 1964.
7. Ченцов Н.Н. Инвариантные функции потерь в задачах математической статистики // Успехи матем. наук. 1966. Т. 22, вып. 1. С. 178-180.
8. Санов И.Н. О вероятности больших уклонений случайных величин // Матем. сб. 1957. Т. 42, № 1. С. 11-44.
9. Ченцов Н.Н. Геометрия "многообразия" распределений вероятностей // Докл. АН СССР. 1964. Т. 158, № 3. С. 543-546.

## АЛГЕБРАИЧЕСКИЕ ОСНОВАНИЯ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ \*)

Общая концепция статистического решающего правила лежит в основе всей современной статистической теории. Каждая конкретная статистическая задача есть (см. [1]) задача принятия решения: статистик, обработав материалы наблюдений, должен сделать выводы о наблюдаемом случайном явлении. Остановимся на этом вопросе подробнее. Математическая модель случайного явления обычно описывается следующим образом. Определяется пространство  $\Omega$  элементарных исходов  $\omega$  идеализированного явления (или эксперимента) и выделяется  $\sigma$ -алгебра  $S$  событий (подмножеств пространства  $\Omega$ ). Затем указывается распределение вероятностей  $P(d\omega)$  исходов — неотрицательная нормированная мера на  $(\Omega, S)$ . В задачах статистики модель наблюдаемого явления известна лишь частично. Зная качественное описание  $(\Omega, S)$  явления и располагая одним или несколькими наблюдениями явления (исходами  $\omega^{(1)}, \dots, \omega^{(N)}$ ), статистик должен оценить неизвестный ему вероятностный закон  $P$  распределения исходов или какие-либо его характеристики.

Таким образом, в статистике решают, грубо говоря, обратную задачу теории вероятностей. Чтобы такая задача была корректно поставленной при бесконечном пространстве исходов, необходима дополнительная априорная информация о неизвестной мере  $P$ . Чаще всего считают заранее известным, что  $P$  принадлежит к некоторому не слишком обширному семейству  $\{P_\theta, \theta \in \Theta\}$ . Для полного описания задачи статистического решения надо указать еще измеримое пространство  $(\Delta, B)$  выводов  $\delta$ , к одному из которых мы должны прийти в ходе решения. Оно является дискретным в задачах различения гипотез, областью многомерного пространства в задачах оценки параметра распределения и т.п.

Всякое решающее правило задачи (неважно — детерминированное или рандомизированное) описывается некоторым переходным распределением вероятностей  $\Pi(\omega; d\delta)$  из  $\Omega$  в  $(\Delta, B)$ . Внесение элементов стохастичности в процесс

\*) Препринт ИПМ, 1977. 17 с. (Перевод доклада на конференции, опубликованного в Algebraic Foundation of Mathematical Statistics // Math. Operationsforsch. Statist. Ser. Statistics. 1978. V. 9, № 2. P. 267-276.)

решения не является принципиальным, поскольку наши выводы все равно будут случайными из-за случайности самих исходов  $\omega^{(1)}, \dots, \omega^{(N)}$ .

Если истинное, неизвестное статистику распределение исходов явления есть  $P(d\omega)$ , то при использовании решающего правила  $\Pi$  распределение выводов будет

$$Q(\cdot) = \int_{\Omega} P(d\omega)\Pi(\omega; \cdot), \quad Q = P\Pi. \quad (1)$$

Таким образом, из априорной информации вытекает, что распределение выводов принадлежит семейству  $\{Q_{\theta} = P_{\theta}\Pi, \theta \in \Theta\}$ , а это позволяет провести предварительный теоретический анализ качества правила  $\Pi$ .

Определим теперь категорию CAP. Каждому измеримому пространству  $(\Omega, S)$  мы сопоставим совокупность  $\text{Cap}(\Omega, S)$  всех распределений вероятностей на  $(\Omega, S)$  — один из объектов категории CAP. Морфизмы категории CAP — преобразования одного объекта в другой будут определяться по формуле (1) всевозможными переходными вероятностями. Единичный морфизм каждого объекта описывается переходным распределением  $\Pi_0(\omega, A) = \chi_A(\omega)$ . Композиция переходных вероятностей из  $\Omega'$  в  $(\Omega'', S'')$  и из  $\Omega''$  в  $(\Omega''', S''')$  есть переходная вероятность из  $\Omega'$  в  $(\Omega''', S''')$ :

$$\int_{\Omega''} \Pi'(\omega'; d\omega'')\Pi''(\omega''; A) = \Pi(\omega'; A). \quad (2)$$

Заметим, что композиция (2) решающих правил описывает составное решающее правило, в котором по закону  $\Pi'$  делается сначала предварительный вывод  $\omega''$ , и только по нему по правилу  $\Pi''$  — окончательный. Эта операция композиции ассоциативна:  $\Pi'(\Pi''\Pi''') = (\Pi'\Pi'')\Pi'''$ . Таким образом, аксиомы алгебраической категории выполнены [2, 3].

Кроме морфизмов, определяемых решающими правилами, для объектов категории CAP имеют смысл и другие операции<sup>1)</sup>. В первую очередь следует указать операцию (тензорного) умножения объектов:

$$\begin{aligned} (\text{Cap}(\Omega', S'), \text{Cap}(\Omega'', S'')) &\xrightarrow{\otimes} \text{Cap}(\Omega' \times \Omega'', S' \otimes S''), \\ (P', P'') &\longrightarrow P = P' \otimes P'', \end{aligned}$$

$$P : P(A_1 \times A_2) = P'(A_1) \cdot P''(A_2) \quad \forall A_1 \in S', \quad \forall A_2 \in S'';$$

позволяющую рассматривать два независимых случайных явления как одно составное.

**ФУНДАМЕНТАЛЬНАЯ ЛЕММА.** *Если морфизмы  $\Pi^{(i)} : \text{Cap}(\Omega^{(i)}, S^{(i)}) \rightarrow \text{Cap}(\Delta^{(i)}, B^{(i)})$ , то существует и единственно переходное распределение вероятностей*

$$\Pi : \text{Cap}\left(\times_i \Omega^{(i)}, \otimes_i S^{(i)}\right) \rightarrow \text{Cap}\left(\times_i \Delta^{(i)}, \otimes_i B^{(i)}\right)$$

со свойством

$$\Pi(\omega^{(1)}, \dots, \omega^{(n)}; A^{(1)} \times \dots \times A^{(n)}) = \Pi^{(1)}(\omega^{(1)}, A^{(1)}) \cdot \dots \cdot \Pi^{(n)}(\omega^{(n)}, A^{(n)}).$$

<sup>1)</sup> Ответ на замечание доктора Хёшеля.

Аналогичные утверждения справедливы для марковских переходов  $\Pi^{(i)}$  из одного и того же пространства  $\Omega$ . Таким образом, категорию CAP можно рассматривать как категорию с умножением.

Выше мы указывали, что категорная алгебра решающих правил образует основу теории статистического вывода. Идя далее, мы можем рассматривать семейства вероятностных законов  $\{P_\theta, \theta \in \Theta\}$  как запараметризованные фигуры и изучать возникающую геометрию фигур с указанной категорией отображений. Эта геометрия является естественным обобщением групповых однородных геометрий в смысле Ф. Клейна.

**ТЕОРЕМА 1.** *Два семейства  $\{P_\theta^{(1)}, \theta \in \Theta\}$  и  $\{P_\theta^{(2)}, \theta \in \Theta\}$  с общим пространством  $\Theta$  параметров эквивалентны тогда и только тогда, когда они конгруэнтны в категорной геометрии, т.е. когда существуют такие морфизмы  $\Pi^{12}$  и  $\Pi^{21}$ , что*

$$P_\theta^1 \Pi^{12} = P_\theta^{(2)}, \quad P_\theta^2 \Pi^{21} = P_\theta^{(1)} \quad \forall \theta \in \Theta. \quad (3)$$

Эта теорема восходит к Д. Блеквеллу [4] (см. также библиографию в [5, 6]). Она утверждает, что класс всех семейств распределений выводов для одного семейства распределения исходов совпадает с аналогичным классом для другого, ему конгруэнтного:  $P_\theta^1 \Pi = P_\theta^{(2)} \Pi^{21} \Pi$ , и наоборот. Вся дополнительная априорная информация в статистической задаче может быть аналогично интерпретирована на языке категорной геометрии [6]<sup>2)</sup>.

Возвратимся к обработке данных. Интуитивно ясно, что каждая выборка содержит какое-то количество полезной информации, причем:

А) при обработке информация может только теряться (монотонность);

Б) информация, доставляемая независимыми источниками, например, независимыми выборками, суммируется (аддитивность).

Таким образом, если ввести информативность<sup>3)</sup> эксперимента, как среднее количество полезной информации в наблюдении, то для информативности выполнены аксиомы А и Б. И хотя понятие информации остается интуитивным, в некоторых статистических задачах можно указать величину “I”, которая естественно выступает в качестве информативности. Такая величина удовлетворяет аксиомам А и Б и описывает асимптотику средней точности выводов в задаче при росте числа наблюдений. Она может быть как числовой, так и матричной величиной. Важнейший пример такой величины — информация по Фишеру.

Необходимо подчеркнуть, что связь между информативностью и средней точностью выводов не сводится к простой пропорциональности. Более того, для разных типов статистических задач эта связь (как и само понятие точности) оказывается различной. Так, в задачах оценки параметра средняя ошибка обратна корню квадратному из информативности. В задачах проверки гипотез вероятность ошибки зависит от информативности экспоненциальным образом

<sup>2)</sup> Категорно-геометрический язык, которому посвящен доклад, не исчерпывает содержания математической статистики, подобно тому как язык теории меры не исчерпывает содержания предельных теорем теории вероятностей. (Ответ на вопрос доктора Нуссбаума о связи категорного подхода с теорией непрерывности Ле Кама [5]).

<sup>3)</sup> Английский перевод русского термина дан профессором М. Стоном и профессором Б. Пенковым.

и т.д. Именно ввиду подобного характера связи при статистической обработке последующие наблюдения приносят, как известно, все меньшую пользу по сравнению с первоначальными<sup>4)</sup>.

Информативность может служить примером аддитивного монотонного инварианта категорий геометрии. Аксиома А описывает монотонное невозрастание относительно категорных отображений. Аксиома Б описывает свойства аддитивности при умножении объектов. Инвариантные и эквивариантные характеристики фигур категорной геометрии суть естественные понятия теории статистического вывода, позволяющие сформулировать ее общие закономерности. Сейчас мы рассмотрим следующую геометрическую задачу. Пусть  $P, Q \in \text{Car}(\Omega, S)$  взаимно абсолютно непрерывны.

Какова инвариантная числовая функция  $\varphi(P, Q)$ , естественно описывающая их непохожесть? И каково ковариантное с  $P$  и  $Q$  распределение  $R$ , лежащее точно посередине  $P$  и  $Q$ ? Опишем, следуя [6], ответы на эти вопросы.

**ТЕОРЕМА 2.** *Если инвариантная в категорной геометрии метрика  $\rho$  задается некоторой нормой,  $\rho(P, Q) = \|P - Q\|$ , то эта метрика отличается от полной вариации разности только на постоянный множитель.*

Расстояние по вариации хорошо известно, но как характеристика непохожести не слишком интересно для статистика. Исследуем теперь вопрос о расстоянии в какой-либо римановой метрике.

**ТЕОРЕМА 3.** *Существует единственная с точностью до множителя инвариантная в категорной геометрии риманова метрика. На гладких семействах она задается квадратичной дифференциальной формой Фишера*

$$ds^2 = \sum_{\alpha, \beta=1}^m d\theta_\alpha d\theta_\beta \cdot \mathbf{M}_p \frac{\partial \ln p(\omega; \theta)}{\partial \theta_\alpha} \frac{\partial \ln p(\omega; \theta)}{\partial \theta_\beta}. \quad (4)$$

В частности, для распределений на конечных множествах  $\Omega = \{\omega_1, \dots, \omega_n\}$  форма Фишера имеет вид

$$ds^2 = \sum_{j=1}^n \frac{(dp_j)^2}{p_j}, \quad p_j := P(\omega_j) \quad \forall j. \quad (5)$$

Расстояние между  $P$  и  $Q$  определяется формулой Бхаттачарья

$$s(P, Q) = 2 \arccos \left( \sum_k \sqrt{p_k q_k} \right), \quad (6)$$

а середина  $R$  дуги между ними — вероятностями

$$r_j = \left[ 2 + 2 \sum_k \sqrt{p_k q_k} \right]^{-1} (p_j + 2\sqrt{p_j q_j} + q_j). \quad (7)$$

Инвариантный тензор Фишера играет фундаментальную роль в математической статистике: через него выписывается неравенство информации [7]. Но с

<sup>4)</sup> Ответ на замечание профессора М. Стона.

соответствующим расстоянием (6) мы в статистике не встречаемся (поскольку оно не удовлетворяет аксиоме Б).

Попробуем обобщить задачу. Будем отыскивать инвариантные линейные связности и их геодезические. Последние полностью определяются правилом построения середины  $R$  отрезка  $PQ$  геодезической. Показано, что инвариантных линейных связностей существует много. В простейшей  $R = P/2 + Q/2$ , и она отвечает измерению расстояния по вариации. Инвариантная риманова связность определяется (7). Но в самой замечательной связности середина  $R$  определяется средним геометрическим мер  $P$  и  $Q$ :

$$r_j = \left[ \sum_k \sqrt{p_k q_k} \right]^{-1} \sqrt{p_j q_j} \quad \forall j. \quad (8)$$

При этом геодезические линии являются экспонентными семействами. Связность (8) выделяется следующим свойством: если семейство мер  $P_s(\cdot)$  экспонентно, то и семейство условных распределений  $P_s(\cdot|H)$  также экспонентное. Но способа мерить похожесть и эта связность непосредственно не дает.

Попытаемся проанализировать причину своей неудачи. Все дело в том, что похожесть случайных явлений несимметрична. Поясним это на грубом примере. Пусть мы хотим установить по результатам наблюдений, не является ли данная девушка ведьмой. Если девушка летает ночью на помеле, то она ведьма; нормальные девушки летать на помеле не умеют. Застав данную девушку в полете, мы можем уверенно сказать, что она — ведьма. Но если за весь период наблюдений этого ни разу не случится, мы не можем наверняка откинуть гипотезу о ее связи с нечистой силой (возможно, в дни наблюдений была нелетная погода, например). Разумеется, Вы можете жениться на этой девушке и точно проверить, есть ли у нее хвостик. Но такой прием выходит за рамки статистического подхода.

Совершенно аналогично, пусть по результатам наблюдений мы хотим узнать, содержатся ли в урне черные шары. Если при Вас хоть единожды вынут из урны черный шар, Вы придете к безошибочному выводу. Если же все время из урны вынимают белые шары, то остается сомнение, не получилась ли эта серия шаров в результате игры случая. Конечно, можно разбить урну и установить ее состав. Но этот прием также выходит за рамки статистического подхода.

Формально эта несимметрия проявляется в следующем. Если конкурируют две простые гипотезы  $P_0$  и  $P_1$ , то (при фиксированной вероятности  $\beta_N$  ошибки второго рода) вероятность  $\alpha_N$  ошибки первого рода (когда ошибочно отвергается нулевая гипотеза) для оптимального решающего правила убывает с ростом числа  $N$  наблюдений как  $\alpha_N \asymp \exp[-N \cdot I(P_1 : P_0)]$  (см. [8]).

Как правило,  $I(P_1 : P_0) \neq I(P_0 : P_1)$ . Здесь

$$I(Q : P) = \int_{\Omega} \left[ \frac{dQ}{dP}(\omega) \ln \frac{dQ}{dP}(\omega) \right] P(d\omega) = \int_{\Omega} \left[ \ln \frac{dQ}{dP}(\omega) \right] Q(d\omega) \quad (10)$$

— информационное уклонение Кульбака–Лейблера–Санова, играющее роль информативности в задаче различения гипотез.

Уклонение  $I(Q : P)$  возникает во многих вопросах математической статистики и по своим геометрическим свойствам является несимметричным обобщением половины квадрата евклидова расстояния; см. ниже теорему 6.

Рассмотрим классическую задачу оценки параметра. Пусть дано гладкое семейство  $\{P_\theta, \theta \in \Theta\}$  возможных распределений на  $(\Omega, S)$ . По  $N$  независимым наблюдениям требуется оценить значение  $\theta$  для наблюдаемого  $P \in \{P_\theta\}$ . Здесь  $\Theta \subset \mathbb{R}^m$ . Точность оценки параметра ограничена снизу неравенством информации. Но статистику чаще нужно оценить сам закон  $P_\theta$ , а не его “имя”  $\theta$ . А при гладкой перепараметризации неравенство информации изменится. Поэтому получим сейчас из контравариантных неравенств Крамера–Рао–Дармуа–Фреше инвариантное следствие.

Прежде всего, в качестве инвариантной функции потерь, характеризующей уклонение оценки  $Q$  от закона  $P$ , возьмем  $2I(Q : P_\theta)$ , переходящее в гауссов квадрат длины для нормального семейства с единичной матрицей ковариаций. Далее, сделаем широкий жест и разрешим оценке  $Q$  принимать любое значение из  $\text{Cap}(\Omega, S)$ , а не только из семейства  $\{P_\theta\}$ . Заметим, что иногда такое расширение может оказаться выгодным. В результате мы приходим к функции риска

$$\mathcal{R}_\Pi(P_\theta) = \int_{\Omega^N} \int_{\text{Cap}(\Omega, S)} 2I(Q : P_\theta) \Pi(\omega^{(1)}, \dots, \omega^{(N)}; d[Q(\cdot)]) P_\theta(d\omega^{(1)}) \dots P_\theta(d\omega^{(N)}). \quad (11)$$

Рассмотрим на  $\{P_\theta\}$  инвариантный объем

$$dV = \sqrt{\det \left( \mathbf{M}_p \frac{\partial \ln p(\omega; \theta)}{\partial \theta_\alpha} \frac{\partial \ln p(\omega; \theta)}{\partial \theta_\beta} \right)_{\alpha, \beta=1}^m} d\theta_1 \dots d\theta_m \quad (12)$$

и средний риск по открытой области  $\Theta$

$$\mathfrak{M}\mathcal{R}_\Pi = \frac{1}{V(\Theta)} \int_{\Theta} \mathcal{R}_\Pi(P - \theta) dV(\theta). \quad (13)$$

**ТЕОРЕМА 4.** *Для гладкого семейства  $\{P_\theta, \theta \in \Theta\}$  взаимно абсолютно непрерывных распределений с открытой областью  $\Theta$  изменения параметра при росте числа  $N$  наблюдений*

$$\lim_{N \rightarrow \infty} N \cdot \inf_{\Pi(N)} \mathfrak{M}\mathcal{R}_{\Pi(N)} = \dim \Theta. \quad (14)$$

То, что левая часть асимптотически больше правой, выводится из неравенства информации. Обратное асимптотическое неравенство вытекает из свойств оценки наибольшего правдоподобия.

Простая и красивая формула (14) имеет важные следствия. Она, в частности, показывает, что неравенство Крамера правильно описывает точность статистических оценок, и что в асимптотической теории в смысле Ле Кама [5] все другие аналогичные информационные неравенства с гладкими функциями потерь могут давать только асимптотически пренебрежимые поправки к главному члену.

Принципиально новые неравенства могут появиться только для негладких семейств, например, для семейств сингулярных друг относительно друга распределений вероятностей или для гладких семейств с самопересечением<sup>5)</sup>.

Рассмотрим теперь задачу определения по наблюдениям плотности априори гладкого закона распределения. Эту задачу иногда неправильно относят к непараметрической статистике. На самом деле эта задача определения счетного числа параметров — прямое обобщение только что рассмотренной классической задачи оценки конечного числа параметров. Поясним это на примере; ср. [9]. Пусть заранее известно, что у искомой плотности  $p(\varphi)$  распределения на окружности

$$\|p''\|^2 = (2\pi)^{-1} \int_0^{2\pi} [p''(\varphi)]^2 d\varphi \leq A^2.$$

Тогда нам достаточно найти ее коэффициенты Фурье

$$c_k = (2\pi)^{-1} \int_0^{2\pi} e^{-ik\varphi} p(\varphi) d\varphi,$$

принимаящие значения в эллипсоиде  $\sum_k k^4 c_k^2 \leq A^2$ .

Эллипсоид изменения параметров имеет разные размеры по разным направлениям. Чем больше номер гармоники, тем меньше соответствующая полуось. Благодаря этому он компактен и задача оценки поставлена для него корректно. Напрашивается следующий (один из возможных) способ оценивания: приближенно найти коэффициенты с малыми номерами, а остальные занулить (требование неотрицательности оценки  $p^*(\varphi) = \sum_k c_k^* e^{ik\varphi}$  мы пока пренебрегаем). Как мы знаем, квадрат точности оценки  $n$  параметров не превосходит  $bnN^{-1}$ , а квадрат точности аппроксимации будет  $g(n) = A^2(n/2 + 1)^{-4}$ , когда  $n$  четно. По теореме Пифагора квадрат суммарной ошибки будет не превосходить

$$bnN^{-1} + g(n), \quad (15)$$

что достигает минимума, когда обе величины примерно одинаковы, т.е. при

$$n = \Gamma(N), \quad \text{где} \quad \Gamma\left(\frac{y}{g(y)}\right) \equiv y. \quad (16)$$

Сама функция потерь будет по порядку величины

$$n_{\text{опт}} N^{-1} \asymp N^{-1} \Gamma(N). \quad (17)$$

С другой стороны, повторяя выкладки, приведшие к (14), нетрудно установить, что

$$\sup_P \mathbf{M}_P \sum_k [c_k^* - c_k(P)]^2 \geq \frac{1}{2} \max_n \min \left\{ \frac{n}{N}, \beta g(n) \right\} \asymp N^{-1} \Gamma(N), \quad (18)$$

где  $\beta > 0$  — некоторая константа,  $\beta g(n)$  — радиус максимального  $n$ -мерного диска, лежащего в эллипсоиде и содержащего только неотрицательные функции. Тем самым, оценки снизу и сверху сомкнуты по порядку убывания.

<sup>5)</sup> Последнее утверждение не является математической теоремой (ответ на вопрос академика У. Бунке).

Подобная, но более точная теорема может быть установлена в терминах введенного выше несимметричного информационного расстояния.

**ТЕОРЕМА 5.** Пусть  $\Phi$  — компактное квазиоднородное экспоненциально выпуклое семейство априори возможных распределений, и пусть квадраты  $\delta_n^2$  и  $\rho_n^2$  его  $n$ -мерных информационных поперечников и его  $n$ -мерных информационных радиусов стремятся к нулю по порядку величины как  $g(n)$ , где  $g''(y) > 0$ ,  $g(\infty) = 0$ .

Тогда максимум  $\Phi$  риска  $\mathcal{R}_{\Pi(N)}(P)$ , определенного в (11) с инвариантной функцией потерь  $2I(Q : P)$ , убывает по порядку величины как  $N^{-1}\Gamma(N)$ , где  $\Gamma$  определено в (16).

Доказательство этой теоремы проводится (см. [6]) аналогично проведенным выше рассуждениям. Только расстояния определяются через информационные уклонения, бесконечномерное семейство  $\Phi$  аппроксимируется  $n$ -мерными экспонентами семействами.

Существенным является требование квазиоднородности семейства  $\Phi$ . Это означает, что существует константа  $C = C(\Phi)$  такая, что

$$\mathbf{D}_R \left[ \ln \frac{dP_1}{dP_2}(\omega) \right] \leq C \mathbf{D}_Q \left[ \ln \frac{dP_1}{dP_2}(\omega) \right] \quad \forall P_1, P_2, Q, R \in \Phi. \quad (19)$$

Отсюда, в частности, следует, что “расстояние туда” и “расстояние обратно” на  $\Phi$  отличаются не слишком сильно:

$$I(Q : P) \leq C^2 I(P : Q) \quad \forall P, Q \in \Phi. \quad (20)$$

В задачах оценки неизвестного закона  $P$  распределения, априори принадлежащего экспонентному семейству или компактному экспоненциально выпуклому семейству, при функции потерь  $2I(Q : P)$  оценки  $Q$  со свойством  $Q \in \{P\}$  образуют полный класс. Это вытекает из следующего предложения, являющегося несимметричным аналогом теоремы Пифагора для информационного уклонения, [10, 6].

**ТЕОРЕМА 6.** Пусть вероятностная мера  $R$  доминируется мерами  $\{P_s\}$ , образующими экспонентное семейство. Если распределение  $P_\sigma \in \{P_s\}$  таково, что

$$\int_{\Omega} \left[ \ln \frac{dP_{s'}}{dP_{s''}}(\omega) \right] (R - P_\sigma)(d\omega) = 0, \quad (21)$$

каковы бы ни были  $P_{s'}, P_{s''} \in \{P_s\}$ , то

$$I(R : P_s) = I(R : P_\sigma) + I(P_\sigma : P_s) \quad \forall P_s. \quad (22)$$

**СЛЕДСТВИЕ.** Мера  $P_\sigma$  является  $I$ -ближайшей к  $R$  точкой экспонентного семейства.

В последнее время ряд предложений возникающей несимметричной Пифагоровой информационной геометрии был переоткрыт Чисаром [11].

Вернемся к поставленному в начале доклада вопросу: какой же вероятностный закон  $R(\cdot)$  естественно считать лежащим точно посередине между взаимно абсолютно непрерывными  $P_0$  и  $P_1$ ? Рассмотрим величину

$$J(P_0, P_1) = \min_{R \ll P} \max_{k=0,1} I(R : P_k). \quad (23)$$

Указанный минимум достигается на законе  $P_v$ ,  $0 < v < 1$ , принадлежащем к экспонентному семейству с плотностями  $p(\omega; u)$ :

$$\ln p(\omega; u) = \ln p_0(\omega) + u[\ln p_1(\omega) - \ln p_0(\omega)] - \Psi(u), \quad (24)$$

где  $\exp \Psi(u)$  — нормирующий делитель,

$$\Psi(u) = \ln \int_{\Omega} \exp[u \ln p_1(\omega) + (1-u) \ln p_0(\omega)] \mu(d\omega). \quad (25)$$

Точка  $P_v$  определяет разбиение множества всех вероятностных мер  $R \ll P$  на области притяжения закона  $P_0$  и закона  $P_1$ . Если проекция  $R$  на  $\{P_u\}$  лежит левее  $P_v$ , то

$$I(R : P_0) > I(R : P_1),$$

и наоборот. Именно эта точка определяет асимптотику при минимаксном различении простых гипотез по Чернову [12].

**ТЕОРЕМА 7.** При оптимальном минимаксном тестировании простых гипотез  $P_0$  и  $P_1$  равные между собой вероятности ошибок первого и второго рода убывают как

$$\exp[-NJ(P_0, P_1)],$$

где  $N$  — число независимых наблюдений, а по (23)

$$J(P_0, P_1) = I(P_v : P_0) = I(P_v : P_1).$$

Таким образом, естественной в каком-то смысле серединой точек  $P_0$  и  $P_1$  приходится считать точку  $P_v$ . Мы видим, что природа коварна, но не злонамерена.

Любопытно, что сходные геометрические конструкции возникают при тестировании нескольких гипотез. Было бы очень интересно выяснить, какова асимптотика вероятности ошибки при минимаксном различении сложных компактных гипотез.

Автор считает приятным долгом поблагодарить всех участников дискуссии по докладу, в особенности У. Бунке, Х.-П. Хёшеля, М. Нуссбаума, Б. Пенкова и М. Стона.

#### СПИСОК ЛИТЕРАТУРЫ

1. Wald A. // Ann. Math. Statist. 1939. V. 10. P. 299–326.
2. Ченцов Н.Н. // Докл. АН СССР. 1965. Т. 164. С. 511–514.
3. Morse N., Sacksteder R. // Ann. Math. Statist. 1966. V. 37. P. 203–214.
4. Blackwell D. // Ann. Math. Statist. 1953. V. 24. P. 265–272.
5. Le Cam L. Notes on asymptotic methods in statistical decision theory. — Montreal, 1974.
6. Ченцов Н.Н. Статистические решающие правила и оптимальные выводы. — М., 1972.
7. Cramer H. Mathematical methods of statistics. — Princeton Univ. Press, 1946.
8. Neyman J., Pearson E.S. // Biometrika. 1928. V. 20. P. 175–240.
9. Ченцов Н.Н. // Докл. АН СССР. 1962. Т. 147. С. 45–48.
10. Ченцов Н.Н. // Матем. заметки. 1968. Т. 4, вып. 3. С. 323–332.
11. Csiszar I. // Ann. Probab. 1975. V. 3. P. 146–158.
12. Chernoff H. // Ann. Math. Statist. 1952. V. 23. P. 493–507.

## МАРКОВСКАЯ ИНВАРИАНТНАЯ ГЕОМЕТРИЯ НА МНОГООБРАЗИЯХ СОСТОЯНИЙ<sup>\*</sup>)

1. Совокупность всех вероятностных мер  $P(\cdot)$  на измеримом пространстве  $(\Omega, \mathcal{A})$  элементарных исходов  $\omega$  является выпуклым подмножеством полуупорядоченного линейного пространства всех счетно-аддитивных зарядов  $R$ , т.е. мер ограниченной вариации на  $(\Omega, \mathcal{A})$ . С полуупорядоченностью этого пространства согласована естественная норма — полная вариация заряда

$$|R| = R^+(\Omega) + R^-(\Omega), \quad (1.1)$$

где  $R^+$  и  $R^-$  — соответственно положительная и отрицательная части  $R$ :

$$R^+ A = \sup_{B \subseteq A} R(B), \quad R^-(A) = - \inf_{B \subseteq A} R(B). \quad (1.2)$$

Естественная норма превращает совокупность всех вероятностных мер на  $(\Omega, \mathcal{A})$  в метрическое пространство с метрикой  $v(P, Q) = |P - Q|$ . Однако хотя эта метрика и встречается в ряде вопросов теории вероятностей, в статистических проблемах с этой характеристикой близости двух распределений вероятностей дело иметь почти не приходится. Другой известной характеристикой близости двух вероятностных законов является расстояние Бхаттачария

$$\sigma(P, Q) = 2 \arccos \int_{\Omega} \sqrt{P(d\omega)} \sqrt{Q(d\omega)}. \quad (1.3)$$

Оно является римановой метрикой и задается для дискретного  $\Omega$  дифференциальной квадратичной формой

$$(\delta\omega)^2 = \int_{\Omega} \frac{[(\delta P)(d\omega)]^2}{P(d\omega)}. \quad (1.4)$$

Сама по себе форма (1.4), называемая *информацией по Фишеру*, играет важную роль при формулировке и доказательстве неравенства информации в теории статистических оценок параметра, но про расстояние (1.3) “в большом” этого сказать нельзя.

В формулировку основных закономерностей теории проверки статистических гипотез входит информационное уклонение Кульбака–Лейблера–Санова

$$\mathcal{D}(P : Q) = \int_{\Omega} [\ln P(d\omega) - \ln Q(d\omega)] P(d\omega), \quad (1.5)$$

называемое также *относительной энтропией*. Как мы показали [11, 13], эта величина является своеобразным несимметричным аналогом (половины) квадрата

<sup>\*</sup>) Итоги науки и техники. Сер. Современные проблемы математики. Новейшие достижения. 1990. Т. 36. С. 69–102 (совм. с Е.А. Морозовой).

расстояния между распределениями  $P$  и  $Q$ , и для этого расстояния справедлив несимметричный аналог теоремы Пифагора. Существуют и другие, правда, несколько менее интересные характеристики близости двух распределений, своего рода несимметричные меры отличия одного распределения от другого; см. [13, 18].

Важным свойством перечисленных характеристик близости является их инвариантность относительно отображений совокупности вероятностных мер, индуцируемых обратимыми измеримыми отображениями выборочного пространства  $(\Omega, \mathcal{A})$ . Эту инвариантность нетрудно вывести из формул для  $\nu, \sigma, \mathcal{E}$ . Несколько сложнее проверить, что при необратимых отображениях индуцированные распределения вероятности сближаются в смысле указанных “характеристик непохожести”. Как показано нами в [13], это свойство “монотонной инвариантности” указанных величин имеет место и относительно более широкой системы всех марковских отображений, т.е. аффинных отображений  $\Pi$  вида

$$\Pi : P \rightarrow P\Pi = Q; \quad Q(\cdot) = \int_{\Omega} P(d\omega) \Pi(\omega; \cdot), \tag{1.6}$$

где  $\Pi(\omega; \cdot)$  — переходное распределение вероятностей. Марковские отображения образуют алгебраическую категорию (см. [9, 28], а также [13]). Классом ее объектов можно считать класс всех измеримых пространств  $(\Omega, \mathcal{A})$  или, как будем считать мы, изоморфный ему класс совокупностей всех распределений вероятностей на каждом из  $(\Omega, \mathcal{A})$ .

Содержательный смысл марковских отображений состоит в том, что они описывают статистические решающие правила в смысле Вальда [31]. Оказывается, что два семейства вероятностных мер  $\{P_{\theta}^{(i)}, \theta \in \Theta\}$  соответственно на  $(\Omega^{(i)}, \mathcal{A}^{(i)})$ ,  $i = 1, 2$ , оба с общим параметрическим множеством  $\Theta$ , эквивалентны в теории статистического вывода, когда они конгруэнтны в категории геометрии, т.е. когда существуют такие марковские отображения  $\Pi^{12}$  и  $\Pi^{21}$ , что

$$P_{\theta}^{(1)} \Pi^{12} = P_{\theta}^{(2)}, \quad P_{\theta}^{(2)} \Pi^{21} = P_{\theta}^{(1)} \quad \forall \theta \in \Theta. \tag{1.7}$$

**ОПРЕДЕЛЕНИЕ 1.** Числовую функцию  $f$ , заданную на всех квадратах объектов категории, назовем *инвариантом* (пар), если

$$(P_1, P_2) \sim (Q_1, Q_2) \implies f(P_1, P_2) = f(Q_1, Q_2) \tag{1.8}$$

для всех конгруэнтных пар, и *монотонным инвариантом*, если всегда

$$f(P_1, P_2) \geq f(P_1\Pi, P_2\Pi). \tag{1.9}$$

Очевидно, всякий монотонный инвариант является инвариантом в смысле (1.8).

Еще одной важной операцией является прямое умножение измеримых пространств и отвечающее ему тензорное умножение совокупностей всех вероятностных мер на них. Это умножение функториально относительно марковской категории; см. [15]. Заметим здесь, что относительная энтропия (1.5) аддитивна относительно этого умножения, что выделяет ее, а также форму (1.4), среди других характеристик непохожести.

2. Случайные явления микромира не описываются схемами классической (как теперь говорят, коммутативной) теории вероятностей, потому что логика квантовых событий не является аристотелевой. Наиболее удобным способом задания объекта некоммутативной теории оказалась следующая алгебраическая схема. задается некоторая инъективная алгебра фон Неймана  $\mathfrak{B}$  ограниченных линейных операторов, действующих на гильбертовом пространстве  $\mathcal{H}$ , например, алгебра  $\mathfrak{B}(\mathcal{H})$  всех таких операторов. Алгебра  $\mathfrak{B}$  является некоммутативным (вообще говоря) аналогом классической (коммутативной) алгебры всех ограниченных измеримых функций на пространстве элементарных исходов. Эрмитовы элементы алгебры  $\mathfrak{B}$  называются (ограниченными) наблюдаемыми. Вероятностное состояние объекта задается неотрицательным нормированным нормальным (т.е. ультраслабо непрерывным, или, что то же, монотонно непрерывным) линейным функционалом  $\Phi$  на  $\mathfrak{B}$ ,  $\Phi : \mathfrak{B} \xrightarrow{H} \mathbb{R}$ , — аналогом математического ожидания по вероятностной мере; см. [29, 22]. Совокупность  $\mathfrak{S}(\mathfrak{B})$  всех состояний  $\Phi$  объекта — выпуклое замкнутое множество в сопряженном пространстве  $\mathfrak{B}_*$ ,  $(\mathfrak{B}_*)^* = \mathfrak{B}$ . Эрмитовы идемпотенты  $E$  алгебры  $\mathfrak{B}$  являются ортопроекторами на соответствующие (замкнутые) подпространства  $\mathcal{E} = E(\mathcal{H})$  пространства  $\mathcal{H}$ . По аналогии с коммутативным случаем, где идемпотенты суть индикаторы измеримых множеств, эти подпространства называют событиями (а также “да-нет”-экспериментами). Очевидно, классическая схема является частным случаем некоммутативной, с коммутативной алгеброй  $\mathfrak{B}$  ограниченных измеримых функций на пространстве исходов.

ОПРЕДЕЛЕНИЕ 2. По аналогии с коммутативным случаем естественно назвать *марковским* всякое аффинное отображение  $\Pi$ :

$$\Pi : \mathfrak{S}(\mathfrak{B}_1) \rightarrow \mathfrak{S}(\mathfrak{B}_2).$$

По линейности  $\Pi$  продолжается до монотонно непрерывного линейного отображения сопряженных пространств (зарядов),  $\Pi : \mathfrak{B}_{1*} \rightarrow \mathfrak{B}_{2*}$ . Сопряженное отображение  $\Pi^* : \mathfrak{B}_2^H \rightarrow \mathfrak{B}_1^H$  является  $\mathbb{R}$ -линейным, неотрицательным, нормальным (монотонно непрерывным) и нормированным. Оно однозначно продолжается до  $\mathbb{C}$ -линейного отображения  $\mathfrak{B}_2 \rightarrow \mathfrak{B}_1$ . Для простоты мы не будем различать  $\Pi$  и  $\Pi^*$ ; условимся, что  $\Pi$  пишется справа от состояния и слева от наблюдаемой. Нетрудно усмотреть, что система всех марковских отображений всех совокупностей  $\mathfrak{S}(\mathfrak{B})$  образует алгебраическую категорию в смысле Эйленберга–Маклейна [21]. Во-первых, тождественное отображение каждого объекта, очевидно, марковское. Во-вторых, композиция (последовательное выполнение) двух марковских отображений снова есть марковское отображение. В-третьих, эта операция композиции ассоциативна, поскольку это всегда имеет место для отображений. Ниже мы будем называть ее *широкой марковской категорией*.

Алгебры  $\mathfrak{B}$  и их преддвойственные пространства можно тензорно перемножать. Стайнспринг [30] заметил, что индуцируемое этим умножением тензорное произведение марковских отображений может не быть неотрицательным и, тем самым, марковским. Он выделил подкласс вполне положительных отображений, замкнутый относительно операции тензорного умножения.

ОПРЕДЕЛЕНИЕ 3. Отображение  $\Pi : \mathfrak{B}_2 \rightarrow \mathfrak{B}_1$  называется *вполне положительным* [30], если все отображения  $\Pi \otimes \text{id}_r$  положительны, где  $\text{id}_r$  — тождественное

отображение  $\mathfrak{B}(\mathcal{H}_r)$  в себя,  $\dim \mathcal{H}_r = r$ :

$$\Pi \otimes \text{id}_r \geq 0 \quad \forall r \in \mathbb{N}. \tag{2.1}$$

Вполне положительные марковские отображения также образуют категорию, которую мы далее будем называть *узкой марковской категорией*. В частности, результат взаимодействия квантовой частицы с частицами (случайной) среды описывается вполне положительным марковским отображением. Далее, по известной теореме Наймарка [6] о продолжении любого разложения единицы до ортогонального выводится [30], что все марковские отображения в коммутативную алгебру фон Неймана или же отображения такой алгебры являются вполне положительными. Последние описывают акт (классического) измерения какой-либо числовой характеристики частицы [8].

Таким образом, определение 1, перенесенное в некоммутативную теорию, дает нам две геометрических эквивалентности семейств состояний — широкую и узкую. Кроме того, по схеме определения 1 можно ввести эквивалентность семейств относительно измерений. Нетрудно установить для этих эквивалентностей цепочку импликаций

$$\{\text{узкая}\} \implies \{\text{широкая}\} \implies \{\text{измерительная}\}. \tag{2.2}$$

Справедливо следующее предложение о максимальных марковских конгруэнтных семействах.

**ТЕОРЕМА 2.1.** *Если вполне положительные марковские отображения  $\Pi^{12}$  и  $\Pi^{21}$  устанавливают конгруэнтность семейств*

$$\{\Phi_\theta^{(i)}, \theta \in \Theta\} \subseteq \mathfrak{S}(\mathfrak{B}_1), \quad i = 1, 2,$$

то

$$\exists \mathfrak{B}'_i, \quad i = 1, 2, \quad \mathfrak{S}(\mathfrak{B}'_i) \Pi^{ij} = \mathfrak{S}(\mathfrak{B}'_j) \supseteq \{\Phi_\theta^{(j)}\}, \quad i \neq j = 1, 2, \tag{2.3}$$

где  $\mathfrak{B}'_i \subseteq \mathfrak{B}_i$  — некоторые инъективные подалгебры фон Неймана.

Сходные утверждения справедливы и для широкой марковской конгруэнтности. Доказательство немедленно вытекает из эргодической теоремы для марковских эндоморфизмов [19, 16]; см. также [2, 17].

**3.** В этой работе для удобства мы ограничимся сепарабельными пространствами  $\mathcal{H}$ . Условие инъективности  $\mathfrak{B}$  позволит ограничиться полными алгебрами  $\mathfrak{B}(\mathcal{H})$ ,  $\mathfrak{S}(\mathfrak{B}(\mathcal{H})) = \mathfrak{S}(\mathcal{H})$ , и использовать матричные обозначения для описания и наблюдаемых, и состояний, и марковских отображений [25, 7]. Для этого следует лишь фиксировать какой-либо ортонормированный базис в  $\mathcal{H}$ . Тогда наблюдаемые  $A$  будут описываться эрмитовыми матрицами  $({}^i a)$ , состояния  $\Phi$  — неотрицательными эрмитовыми матрицами  $(\varphi_j^i)$  со следом единица. Чтобы различать эти объекты, в первом случае индексы будут писаться слева. Для описания марковского отображения  $\Pi : \mathfrak{S}(\mathcal{H}^{(1)}) \rightarrow \mathfrak{S}(\mathcal{H}^{(2)})$  надо указать базисы в  $\mathcal{H}^{(1)}$  и  $\mathcal{H}^{(2)}$ . Эти начальный и конечный базисы можно выбрать различными, даже если  $\mathcal{H}^{(1)} = \mathcal{H}^{(2)}$ .

ЛЕММА 3.1. Марковское отображение  $\mathbb{I} : \mathfrak{S}(\mathcal{H}^{(1)}) \rightarrow \mathfrak{S}(\mathcal{H}^{(2)})$  описывается четырехвалентным (унитарным) тензором  $({}^i\omega_l^k)$  со свойствами

$${}^i\omega_l^k = \overline{{}^j\omega_k^l} \quad \forall i, j, k, l, \quad (3.1)$$

$$\mathrm{tr}_{\mathcal{H}^{(2)}} \mathbb{I} := \sum_k {}^i\omega_k^k = {}^i\delta, \quad (3.2)$$

$$\sum_{i,j,k,l} \xi^i \xi_j {}^i\omega_l^k \eta_l \eta \geq 0, \quad (3.3)$$

$$\forall (\bar{\xi}_1, \dots, \bar{\xi}_n, \dots) = (\xi^1, \dots, \xi^n, \dots) \quad \forall ({}^1\bar{\eta}, \dots, {}^n\bar{\eta}, \dots) = ({}^1\eta, \dots, {}^m\eta, \dots).$$

Здесь  ${}^i\delta$  — символ Кронекера, черта означает комплексное сопряжение и подразумевается суммирование по совпадающим верхнему и нижнему или левому и правому индексам. Смысл, в котором понимается суммирование в счетномерном случае, установлен в [7]. Условие (3.1) обеспечивает сохранение эрмитовости, условие (3.3) — сохранение неотрицательности, условие (3.2) — сохранение величины следа. Покажем теперь, как упрощается теория Стайнспринга [30] в конечномерном случае.

ЛЕММА 3.2. Матрица  $({}^i\omega_l^k)$  вполне положительного марковского отображения удовлетворяет условиям (3.1), (3.2) и условию

$$\sum_{i,j,k,l} {}^k\zeta_j^i \zeta_j^i {}^i\omega_l^k \geq 0 \quad (3.4)$$

$$\forall ({}^1\zeta^1, \dots, {}^n\zeta^m, \dots) = ({}^1\bar{\zeta}_1, \dots, {}^n\bar{\zeta}_m, \dots),$$

более ограничительно, чем условие (3.3).

ЛЕММА 3.3. Матрица  $({}^i\omega_l^k)$  вполне положительного марковского отображения допускает ([24]) разложение

$${}^i\omega_l^k = \sum_s {}^i\omega_l^k(s) {}_j\omega_l(s) \quad \forall i, j, k, l, \quad (3.5)$$

$${}^i\omega_l^k(s) = \overline{{}_i\omega_k(s)} \quad \forall i, k, s.$$

СЛЕДСТВИЕ. Композиция вполне положительных марковских отображений вполне положительна.

Доказательство. Из условия (3.4) вытекает, что матрица  $({}^i\omega_l^k)$  может быть интерпретирована как матрица неотрицательного эрмитова оператора, действующего в пространстве  $\mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)}$ , откуда следует (3.5). Далее, любой неотрицательный ядерный эрмитов оператор в пространстве  $\mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)}$  также разлагается в сумму тензоров вида  $b^{ip} b_{jq}$ , а  $\mathbb{I} \otimes \mathrm{id}_r$  — образ каждого такого тензора — неотрицателен в силу

$$\sum b^{ip} \cdot b_{jq} \cdot {}^i\omega_l^k \cdot {}_j\omega_l \cdot {}^p\delta^u \cdot {}_q\delta_v \cdot {}^k u a \cdot {}_l v a = \left| \sum b^{ip} \cdot {}^i\omega_l^k \cdot {}^k p a \right|^2 \geq 0.$$

Наконец, условие (3.4) необходимо для вполне положительности, потому что неотрицательный эрмитов оператор в  $\mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)}$  с матрицей  $(\delta^{ip} \delta_{jq})$  отображается

$\mathcal{H} \otimes \text{id}_r$  в оператор с матрицей  $(f_{iv}^{ku})$ , где  $f_{iv}^{ku} = {}^u u_i^k$  при всех значениях индексов. В заключение отметим, что разложение Крауса справедливо и в счетномерном случае.

4. Взвешенная разность двух состояний одного объекта принадлежит эрмитовой части предсопряженного пространства  $\mathfrak{B}_*^H = \text{Lin } \mathfrak{S}(\mathfrak{B})$ . Она описывается ядерным оператором.

ОПРЕДЕЛЕНИЕ 5. Пусть  $\Gamma \in \mathfrak{B}_*^H$ . Его положительную  $\Gamma^{(+)}$  и отрицательную  $\Gamma^{(-)}$  части определим через спектральное разложение:

$$\Gamma = \sum \gamma_k E_k, \quad \Gamma^{(+)} = \sum_{\gamma > 0} \gamma_k E_k, \quad \Gamma^{(-)} = - \sum_{\gamma < 0} \gamma_k E_k. \quad (4.1)$$

Величины

$$\begin{aligned} \text{tr}^{(+)} \Gamma &= \text{tr}^{(+)} \Gamma^{(+)} = \sum_{\gamma > 0} \gamma_k, \\ \text{tr}^{(-)} \Gamma &= - \sum_{\gamma < 0} \gamma_k, \quad |\Gamma| = \sum |\gamma_k| \end{aligned} \quad (4.2)$$

назовем соответственно *положительной*, *отрицательной* и *полной вариацией*  $\Gamma$ . Разложения  $\Gamma = \Gamma^{(+)} - \Gamma^{(-)}$  и  $\mathcal{H} = \mathcal{E}^{(+)} \oplus \mathcal{E}^{(-)}$ , где  $E^{(+)} = \sum_{\gamma > 0} E_k$ ,  $E^{(-)} = \sum_{\gamma < 0} E_k$ , назовем *расщеплениями*.

ЛЕММА 4.1. *Положительная и отрицательная вариации  $\Gamma \in \mathfrak{B}_*^H$  являются монотонными инвариантами в широкой марковской категории. Каждое  $\Gamma \in \mathfrak{B}_*^H$  конгруентно заряду  $\Gamma_{\text{red}}$  на двугатомном измеримом пространстве  $\Omega = \{\omega_1, \omega_2\}$ :*

$$\Gamma_{\text{red}}(\omega_1) = \text{tr}^{(+)} \Gamma, \quad \Gamma_{\text{red}}(\omega_2) = -\text{tr}^{(-)} \Gamma. \quad (4.3)$$

СЛЕДСТВИЕ. *Положительная и отрицательная вариации  $\Gamma \in \mathfrak{B}_*^H$  образуют полную систему инвариантов  $\Gamma$  в обеих марковских категориях.*

Доказательство. Выбрав за базисные в  $\mathcal{H}^{(1)}$  и  $\mathcal{H}^{(2)}$  собственные векторы  $\Gamma$  и  $\mathcal{H}\mathcal{H}$ , проводим построения, известные в коммутативной теории; см. [13, лемма 5.9].

Применяя эти результаты к линейной комбинации состояний  $\Psi - z\Phi$ , получаем два семейства

$$\mu^{(\pm)}(z) = \text{tr}(\Psi - z\Phi)^{(\pm)}, \quad \mu^{(+)}(z) - \mu^{(-)}(z) = 1 - z \quad \forall z \in \mathbb{R} \quad (4.4)$$

монотонных инвариантов пары  $(\Phi, \Psi)$ . Любое из них задает полную систему инвариантов, когда  $\Phi$  и  $\Psi$  коммутируют как операторы; см. [13].

В общем случае вопрос о полной системе инвариантов пока не решен [3]. Соответствующее наше утверждение в [26] является ошибочным, так как коммутиримость состояний есть инвариантное свойство в узкой марковской категории.

5. Пространства с метрикой должны удовлетворять аксиомам: 1°)  $\rho(x, y) \geq 0$ ; 2°)  $\rho(x, y) = 0 \implies y = x$ ; 3°)  $\rho(x, y) = \rho(y, x)$ ; 4°)  $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$ . Когда выполнены только аксиомы 1°), 3°), 4°), говорят о псевдометрике. Мы рассмотрим более экзотический случай несимметричной метрики, когда не выполнена аксиома 3°) (но аксиома 4°) выполнена именно в указанной форме).

ТЕОРЕМА 5.1. Если на всех объектах  $\mathfrak{S}(\mathfrak{B})$  задана метрика  $\rho(\Phi, \Psi)$ , монотонно инвариантная в узкой марковской категории, то

$$\rho(\Phi, \Psi) \geq \frac{1}{8} \rho(R_{1/2}, R_{1/4}) \cdot |\Phi - \Psi|, \quad (5.1)$$

где  $|\Phi - \Psi|$  — полная вариация (следовая норма) разности  $\Phi - \Psi$ ,  $R_\theta$  — распределение вероятности на алгебре с двумя атомами,

$$R_\theta : \{R_\theta(\omega_1) = \theta, \quad R_\theta(\omega_2) = 1 - \theta\}. \quad (5.2)$$

ЗАМЕЧАНИЕ. Если метрика  $\rho$  несимметрична, то

$$\rho(\Phi, \Psi) \geq \frac{1}{8} |\Phi - \Psi| \cdot \min \{ \rho(R_{1/2}, R_{1/4}), \rho(R_{1/4}, R_{1/2}) \}.$$

Доказательство. Пусть  $\mathcal{H} = \mathcal{E}^{(+)} \otimes \mathcal{E}^{(-)}$  — расщепление для ядерного оператора  $\Phi - \Psi$ . Сужение состояний на двухатомную булеву алгебру с атомами  $\mathcal{E}^{(+)}$  и  $\mathcal{E}^{(-)}$  леммы 4.1 — вполне положительное марковское отображение. Отсюда

$$\rho(\Phi, \Psi) \geq \rho(\Phi_{\text{red}}, \Psi_{\text{red}}), \quad |\Phi - \Psi| = |\Phi_{\text{red}} - \Psi_{\text{red}}|.$$

Следовательно, достаточно доказать теорему для распределений вероятностей на двухатомной булевой алгебре. Каждое распределение вероятностей  $R_\theta$  на ней описывается (5.2) вероятностью первого атома  $R_\theta(\omega_1) = \theta$ ,  $0 \leq \theta \leq 1$ . Марковский эндоморфизм совокупности  $\{R_\theta, 0 \leq \theta \leq 1\}$  описывается аффинным отображением отрезка  $\{\theta : 0 \leq \theta \leq 1\}$  в себя. При этом

$$|R_{\theta'} - R_{\theta''}| = 2|\theta' - \theta''|. \quad (5.3)$$

ЛЕММА 5.2. Если  $0 \leq \theta'_1 \leq \theta''_1 \leq \theta''_2 \leq \theta'_2 \leq 1$ , то марковский эндоморфизм  $\Pi$  совокупности  $\{R_\theta\}$  с матрицей

$$\left( \begin{array}{cc} \frac{(1 - \theta'_1)(1 - \theta''_2) - (1 - \theta'_2)(1 - \theta''_1)}{\theta'_2 - \theta'_1} & \frac{(1 - \theta'_1)\theta''_2 - (1 - \theta'_2)\theta''_1}{\theta'_2 - \theta'_1} \\ \frac{\theta'_2(1 - \theta''_1) - \theta'_1(1 - \theta''_2)}{\theta'_2 - \theta'_1} & \frac{\theta'_2\theta''_1 - \theta'_1\theta''_2}{\theta'_2 - \theta'_1} \end{array} \right) \quad (5.4)$$

отображает  $R_{\theta'_1}$  в  $R_{\theta''_1}$ , а  $R_{\theta'_2}$  в  $R_{\theta''_2}$ .

СЛЕДСТВИЕ. Если  $f(\Phi, \Psi)$  — монотонный инвариант узкой марковской категории, то  $f(R_{\theta'_1}, R_{\theta'_2}) \geq f(R_{\theta''_1}, R_{\theta''_2})$  в условиях леммы.

ЛЕММА 5.3. Пусть  $0 \leq \theta_0 \leq 1$ ,  $0 \leq q \leq 1$ ,  $\theta_k = \theta_0 q^k$ . Тогда марковские эндоморфизмы совокупности  $\{R_\theta\}$  с матрицами

$$\mathbb{I}^n = \begin{pmatrix} q^n & 1 - q^n \\ 0 & 1 \end{pmatrix} \quad (5.5)$$

отображают по правилу  $R_{\theta_k} \mathbb{I}^n = R_{\theta_{k+n}}$ . При этом

$$|R_{\theta_{k-1}} - R_{\theta_k}| = q^{k-1} |R_{\theta_0} - R_{\theta_1}|. \quad (5.6)$$

Обе леммы доказываются простыми вычислениями. Перейдем к доказательству теоремы. Пусть  $\Phi_{\text{red}} = \langle\langle p_1, 1 - p_1 \rangle\rangle$ ,  $\Psi_{\text{red}} = \langle\langle p_2, 1 - p_2 \rangle\rangle$ ,  $\theta_0 = \max\{p_1, p_2, 1 - p_1, 1 - p_2\}$ . Не ограничивая общности, можно считать  $\theta_0 = p_1$ . В противном случае перенумеруем исходы (что описывается обратимым марковским отображением распределений) или переставим аргументы. Заметим, сверх того, что при  $p_1 = p_2$  неравенство тривиально выполнено. Поэтому мы можем считать, что  $\theta_0 = p_1 > p_2$  и  $\theta_0 = p_1 > 1/2$ . Обозначим  $q = p_2 \cdot p_1^{-1} < 1$  и построим точки  $\theta_k = \theta_0 \cdot q^k$ ,  $k = 1, \dots, \nu$ , где  $\theta_{\nu-1} > 1/4$ ,  $\theta_\nu < 1/4$ , так что  $q^k = \theta_k \cdot \theta_0^{-1} > 1/4$ ,  $k = 0, 1, \dots, \nu - 1$ . По монотонности  $\rho$ , неравенству треугольника, следствию из леммы 5.2 и оценке (5.6)

$$\nu \cdot \rho(\Phi_{\text{red}}, \Psi_{\text{red}}) \geq \rho(R_{\theta_0}, R_{\theta_1}) + \dots + \rho(R_{\theta_{\nu-1}}, R_{\theta_\nu}) \geq \rho(R_{\theta_0}, R_{\theta_\nu}) \geq \rho(R_{1/2}, R_{1/4}),$$

$$\nu \cdot |\Phi_{\text{red}} - \Psi_{\text{red}}| = 2|\theta_0 - \theta_1| + \dots + 2q^{-\nu+1}|\theta_{\nu-1} - \theta_\nu| \leq 8|\theta_0 - \theta_\nu| \leq 8.$$

Поделив первое из неравенств на второе, получаем (5.1) для двухатомных распределений вероятностей, что и требовалось. Для несимметричных расстояний может быть  $p_2 > p_1$ . Тогда надо строить точки  $\theta_k = \theta_0 \cdot q^k$ , где  $\theta_0 = p_2 \cdot q = p_1 \cdot p_2^{-1}$ .

**6.** Изучим строение естественных функций потерь  $L(\Phi, \Psi)$  в задаче оценки состояния. Мы предполагаем, что функция  $L$  задана на всех квадратах объектов  $\mathfrak{S}(\mathfrak{B})$ , неотрицательна:

$$L(\Phi, \Phi) = 0; \quad \Phi \neq \Psi \implies L(\Phi, \Psi) > 0, \tag{6.1}$$

и монотонна в узкой категории:

$$L(\Phi\Pi, \Psi\Pi) \leq L(\Phi, \Psi). \tag{6.2}$$

Мы допускаем, что  $L$  может принимать значение  $+\infty$ .

Рассмотрим сперва значения  $l(x, y) = L(R_x, R_y)$  на двухатомной булевой алгебре.

**ЛЕММА 6.1.** *Функция  $l(x, y)$  на квадрате  $0 \leq x, y \leq 1$  тождественно равна нулю на диагонали  $x = y$  и строго положительна вне нее. Функция*

$$f(z) = \inf_{|x-y| \geq z} l(x, y) \tag{6.3}$$

*монотонна и строго положительна при  $z > 0$ ,  $f(0) = 0$ .*

**Доказательство.** Первые два утверждения следуют из (6.1), а монотонность  $f$  следует из определения (6.3). Рассмотрим подмножество  $Q(z) = \{(x, y) : |x - y| \geq z, 0 \leq x, y \leq 1\}$ . Оно компактно. Поэтому существует сходящаяся последовательность

$$(x_n, y_n) \rightarrow (x^*, y^*) \in Q(z), \quad \lim l(x_n, y_n) = f(z).$$

Пусть  $x^* < y^*$ . Возьмем любой отрезок  $[x_0, y_0]$  такой, что  $x^* < x_0 < y_0 < y^*$ . Тогда  $[x_0, y_0] \subset [x_n, y_n]$  при  $n \geq n_0$ . Отсюда  $l(x_0, y_0) \leq l(x_n, y_n)$  при  $n \geq n_0$  и  $0 < l(x_0, y_0) \leq f(z)$ . Если  $x^* > y^*$ , проведенное рассуждение надо применить к отрезкам вида  $[y, x]$ .

ТЕОРЕМА 6.2. Если заданная на квадратах объектов  $\mathfrak{S}(\mathfrak{B})$  функция  $L$  удовлетворяет условиям (6.1) и (6.2), то

$$L(\Phi, \Psi) \geq f(|\Phi - \Psi|), \quad (6.4)$$

где функция  $f(z)$ , определенная формулой (6.3), монотонна и строго положительна при  $z > 0$ ,  $f(0) = 0$ .

Доказательство. По монотонности  $L$  и лемме 4.1, примененной к разности  $(\Phi - \Psi) \in \mathfrak{B}_*^H$ ,

$$L(\Phi, \Psi) \geq L(\Phi_{\text{red}}, \Psi_{\text{red}}) \geq f(|\Phi_{\text{red}} - \Psi_{\text{red}}|) = f(|\Phi - \Psi|),$$

где среднее неравенство вытекает из леммы 6.1.

Если монотонная функция  $F(\Phi, \Psi)$  вместо (6.1) удовлетворяет условию  $F(\Phi, \Psi) \neq F(\Phi, \Phi)$  при  $\Phi \neq \Psi$ , то  $F(\Phi, \Psi) \geq c + f(|\Phi - \Psi|)$ , где  $c = F(\Phi, \Phi) = F(\Psi, \Psi)$ , а функция  $f$  построена по (6.3) для  $L(\Phi, \Psi) = F(\Phi, \Psi) - c$ .

7. Инвариантные римановы метрики особенно интересны тем, что определяющие их эквивариантные дифференциальные квадратичные формы позволяют выписывать информационные неравенства; см. [8].

Фиксируем какой-либо собственный базис ядерного оператора состояния  $\Phi$ ,  $\mathcal{H} = \oplus \mathcal{L}^{(k)}$ . В нем  $\Phi$  будет описываться диагональной матрицей  $\text{diag}(p_1, p_2, \dots)$ , где  $p_j = \varphi_j^j$  — собственные числа. В этом же базисе дифференциал  $d\Phi$  будет описываться эрмитовой матрицей со следом нуль:

$$d\varphi_l^k = \xi_l^k + i\eta_l^k, \quad (7.1)$$

где  $\xi$  и  $\eta$  — вещественные и мнимые части дифференциалов, а символ  $d$  опущен для упрощения обозначений. Имеют место связи

$$\xi_l^k = \xi_k^l, \quad \eta_l^k = -\eta_k^l \quad \forall k, l; \quad \eta_k^k = 0 \quad \forall k; \quad \sum_k \xi_k^k = 0. \quad (7.2)$$

Таким образом, полную систему линейных вещественных координат в касательном к  $\mathfrak{S}$  пространстве образуют переменные  $\xi_l^k, \eta_l^k$  при  $k < l$ , а также все переменные  $\xi_k^k$ , кроме одной из них. Для упрощения удобно записать искомую квадратичную форму через все  $\xi_k^k$  (см. [28]) с помощью коэффициентов  $u, v, \omega$  полярной билинейной формы:

$$Q_\Phi(d\Phi) = \sum u_{ji}^{ik} \xi_j^i \xi_l^k + 2 \sum v_{ji}^{ik} \xi_j^i \eta_l^k + \sum \omega_{ji}^{ik} \eta_j^i \eta_l^k. \quad (7.3)$$

Простейшие вполне положительные марковские отображения действуют по формуле

$$\Phi \Pi_s = S \Phi S^*, \quad ({}_j^i w_l^k) = ({}^i s^k \cdot {}_j s_l), \quad (7.4)$$

где  $S$  — унитарный оператор. Отображение  $\Pi_s$ , очевидно, обратимо, и  $\{\Phi, d\Phi\} \sim \{\Phi \Pi_s, d\Phi \Pi_s\}$ .

ЛЕММА 7.1. Если дифференциальная квадратичная форма (7.3) унитарно инвариантна  $Q_\Phi(d\Phi) = Q_{U\Phi U^*}(dU\Phi U^*)$ , то она тождественно равна

$$\sum_j u_{jj}^{jj} \xi_j^j \xi_j^j + \sum_{j < k} (u_{jk}^{jk} + u_{kj}^{kj}) \xi_j^j \xi_k^k + 2 \sum_{j > k} u_{kk}^{jj} [(\xi_k^j)^2 + (\eta_k^j)^2]. \quad (7.5)$$

Доказательство. Пусть симметрия  $S$  есть отражение первой оси  $\mathcal{L}^{(1)}$ . Оно меняет знак у элементов первой строки и первого столбца, кроме диагонального. Так как  $\{\Phi, d\Phi\} \sim \{\Phi III_s, d\Phi III_s\}$ , то значение формы  $Q$  сохраняется, и оно тождественно равно полусумме исходного  $Q$  и выражения, отличающегося знаком минус при членах, в которые индекс 1 входит один или три раза. В полусумме эти члены взаимно уничтожаются, а остальные сохраняются. Прделаем эту операцию последовательно при всех  $k = 1, 2, \dots$ . В результате придем к форме

$$Q_\Phi(d\Phi) = \sum u_{jj}^{jj} (\xi_j^j)^2 + \sum (u_{jk}^{jk} + u_{kj}^{kj}) \xi_j^j \xi_k^k + \sum u_{kk}^{jj} (\xi_k^j)^2 + 2 \sum v_{kk}^{jj} \xi_k^j \eta_k^j + \sum w_{kk}^{jj} (\eta_k^j)^2, \quad (7.6)$$

где все двойные суммы взяты по парам  $j < k$ .

Теперь применим к аргументам  $\Phi$  и  $d\Phi$  операцию поворота первой “оси” в себе, индуцированного умножением на мнимую единицу. При этом  $\xi_k^1 \rightarrow \eta_k^1, \eta_k^1 \rightarrow -\xi_k^1$ . Снова усредняя полученные результаты, уничтожаем члены при  $\xi_k^1 \eta_k^1$  и объединяем слагаемые  $(\xi_k^1)^2 + (\eta_k^1)^2$ . Если в исходной форме положить все  $\eta_k^1 = 0$ , то из  $\Pi$ -инвариантности формы  $Q$  получаем  $u_{kk}^{11} = w_{kk}^{11}$ . Прделаем повороты при всех  $j = 1, 2, \dots$ , приходим от (7.6) к (7.5).

ТЕОРЕМА 7.2. Если дифференциальная квадратичная форма (7.3) инвариантна в узкой марковской категории, то она тождественно равна

$$Q_\Phi(d\Phi) = c \cdot \sum_k (p_k)^{-1} (\xi_k^k)^2 + 2 \sum_{j < k} c(p_j, p_k) |d\varphi_k^j|^2, \quad (7.7)$$

где  $c(x, y)$  — функция в области  $0 \leq x, y \leq x + y \leq 1$ , причем  $c(x, x) = c \cdot x^{-1}$ ,  $c(x, y) = c(y, x)$ .

Доказательство. В коммутативной теории инвариантная риманова метрика единственна с точностью до множителя. Отсюда по (1.10)  $u_{jk}^{jk} = 0$  при  $j \neq k$ ,  $u_{kk}^{kk} = c(p_k)^{-1}$ , и первые два слагаемых (7.5) дают первое слагаемое (7.7). Построим теперь марковское отображение  $III$  из  $\mathfrak{S}(\mathfrak{B})$  на  $\mathfrak{S}_{2,1} = \mathfrak{S}(\mathfrak{B}(\mathcal{L}^{(j)} \oplus \mathcal{L}^{(k)}) \oplus \mathbb{C})$  по формуле  $\Psi III = \Gamma \oplus \gamma$ ,

$$\Gamma = (E_j + E_k) \Psi (E_j + E_k), \quad \gamma = \text{tr } \Psi - \psi_j^j - \psi_k^k, \quad (7.8)$$

и отображение  $\mathcal{J} : \mathfrak{S}_{2,1} \rightarrow \mathfrak{S}(\mathfrak{B})$ ,

$$(\Gamma \oplus \gamma) \mathcal{J} = \Gamma \oplus \gamma (H - E_j - E_k) \Phi (H - E_j - E_k). \quad (7.9)$$

Легко проверить по критерию (3.4) их вполне положительность. На операторе  $\Psi = \Phi$  и дифференциале  $d\Psi = \xi_k^j$  отображения  $III$  и  $\mathcal{J}$  взаимно обратны,  $\gamma(\Phi) = 1 - p_j - p_k$ . Отсюда  $u_{kk}^{jj}(\Phi) = u_{kk}^{jj}(\Phi III) = u_{kk}^{jj}(p_j, p_k)$ . Поскольку перенумерация осей  $\mathcal{L}^{(k)}$  есть унитарное преобразование, то  $u_{kk}^{jj}(x, y) = c(x, y)$ .

Если  $p_j = p_k$ , то любой поворот плоскости  $\mathcal{L}^{(j)} \oplus \mathcal{L}^{(k)}$  оставляет  $\Phi$  инвариантным. Поворот  $\Pi$  на  $\pi/4$  переводит  $d\Phi = (\xi_k^j, \eta_k^j)$  в  $d\Phi\Pi = (\gamma_j^j, \gamma_k^k)$ ,  $\gamma_j^j = |\xi_k^j + i\eta_k^j| = -\gamma_k^k$ , что дает  $c(p, p) = c/p$ .

ЛЕММА 7.3. Если инвариантная дифференциальная квадратичная форма  $Q_\Phi$  слабо непрерывна по  $\Phi$ , то функция  $c(x, y)$  в ее разложении (7.7) положительно однородна в минус первой степени:

$$c(x, y) = \lambda \cdot c(\lambda x, \lambda y) \quad \forall \lambda: 0 < \lambda \leq (x + y)^{-1}. \quad (7.10)$$

Доказательство. Рассмотрим марковское отображение разложения  $\Pi: \mathfrak{S}(\mathfrak{B}) \rightarrow \mathfrak{S}(\mathfrak{B} \otimes \mathfrak{B}(\mathcal{H}_n))$ , действующее по формуле

$$\Psi \rightarrow \Psi \otimes \frac{1}{n} \mathbf{I}_n. \quad (7.11)$$

Оно обратимо на образе  $\mathfrak{S}(\mathfrak{B})\Pi$ . По конгруэнтности  $\{\Phi, d\Phi\}$  и его образа

$$c(x, y) = n \cdot n^{-2} c(xn^{-1}, yn^{-1}) = \nu^{-1} c(x\nu^{-1}, y\nu^{-1}).$$

Обозначая  $x' = xn^{-1}$ ,  $y' = yn^{-1}$ , приходим к (7.10) с рациональным  $\lambda$ . Справедливость в общем случае следует из непрерывности.

ТЕОРЕМА 7.4. Всякая инвариантная дифференциальная квадратичная форма со свойством (7.10), в частности, слабо непрерывна, удовлетворяет принципу сложения информации

$$Q_{\Phi \otimes \Psi}(d[\Phi \otimes \Psi]) = Q_\Phi(d\Phi) + Q_\Psi(d\Psi). \quad (7.12)$$

Доказательство. Выберем собственные базисы операторов  $\Phi = \text{diag}\langle p_1, p_2, \dots \rangle$  и  $\Psi = \text{diag}\langle q_1, q_2, \dots \rangle$ . В тензорном произведении этих базисов только матрица дифференциала  $d(\Phi \otimes \Psi) = \Phi \otimes (d\Psi) + (d\Phi) \otimes \Psi$  имеет много нулей, и

$$\begin{aligned} Q(d[\Phi \otimes \Psi]) &= \sum_{i,k} (p_i q_k)^{-1} [p_i d\psi_k^k + q_k d\varphi_i^i]^2 + \\ &+ 2 \sum_{i,k < l} c(p_i q_k, p_i q_l) |p_i d\psi_l^k|^2 + 2 \sum_{i < j; k} c(p_i q_k, p_j q_k) |q_k d\varphi_j^i|^2 = \\ &= \sum_k (q_k)^{-1} |d\psi_k^k|^2 + \sum_i (p_i)^{-1} |d\varphi_i^i|^2 + 2 \sum_{k < l} c(q_k, q_l) |d\psi_l^k|^2 + 2 \sum_{i < j} c(p_i, p_j) |d\varphi_j^i|^2, \end{aligned}$$

где мы воспользовались правилом (7.10) и равенствами

$$\sum_i p_i = \sum_k q_k = 1, \quad \sum_i d\varphi_i^i = \sum_k d\varphi_k^k = 0.$$

8. Если предположить, что риманова метрика с формой  $Q$  не только инвариантна, но и монотонна в категории, для коэффициента  $c(x, y)$  формы (7.7) можно указать оценки.

ЛЕММА 8.1. Если метрика монотонна, то

$$c(x, y) \geq \theta^2 \cdot c(\theta x, \theta y) \quad \forall \theta: 0 \leq \theta \leq 1. \quad (8.1)$$

СЛЕДСТВИЕ. В условиях леммы свойство (7.10) выполняется при всех  $\lambda$  без требования слабой непрерывности.

**Доказательство.** Зададим вполне положительный марковский эндоморфизм  $\mathbb{I}$  совокупности  $\mathfrak{S}_{2,1} = \mathfrak{S}(\mathfrak{B}(\mathcal{H}_2) \oplus \mathbb{C})$  состояний  $\Psi = P \oplus q$ ,  $q + \text{tr } P = 1$  формулой

$$\Psi\mathbb{I} = \theta P \oplus [q + (1 - \theta)\text{tr } P]. \quad (8.2)$$

Его можно продолжить до эндоморфизма совокупности  $\mathfrak{S}(\mathcal{H}_3) \supset \mathfrak{S}_{2,1}$  описываемого в базисе  $\mathcal{H}_2 \oplus \mathbb{C} = \mathcal{H}_3$  четыре-матрицей третьего порядка  ${}^i\mathbb{I}_j^k$

$$\left( \begin{array}{c|c|c} \begin{array}{c} {}^1\mathbb{I}_1 \\ \hline {}^2\mathbb{I}_1 \\ \hline {}^3\mathbb{I}_1 \end{array} & \begin{array}{c} {}^1\mathbb{I}_2 \\ \hline {}^2\mathbb{I}_2 \\ \hline {}^3\mathbb{I}_2 \end{array} & \begin{array}{c} {}^1\mathbb{I}_3 \\ \hline {}^2\mathbb{I}_3 \\ \hline {}^3\mathbb{I}_3 \end{array} \\ \hline \begin{array}{c} 0 \\ \hline 0 \\ \hline 0 \end{array} & \begin{array}{c} 0 \\ \hline 0 \\ \hline 0 \end{array} & \begin{array}{c} 0 \\ \hline 0 \\ \hline 1 \end{array} \end{array} \right) = \left( \begin{array}{ccc|ccc|ccc} \theta & 0 & 0 & 0 & \theta & 0 & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & & & 0 \\ 0 & 0 & 1 - \theta & 0 & 0 & 0 & & & \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & & & 0 \\ \theta & 0 & 0 & 0 & \theta & 0 & & & \\ 0 & 0 & 0 & 0 & 0 & 1 - \theta & & & \\ \hline & & & 0 & & & 0 & 0 & 0 \\ & & & 0 & & & 0 & 0 & 0 \\ & & & 0 & & & 0 & 0 & 1 \end{array} \right).$$

Ее вполне положительность вытекает из критерия (3.4). Если  $\Phi = \text{diag} \langle x, y, 1 - x - y \rangle$ , то  $\Phi\mathbb{I} = \text{diag} \langle \theta x, \theta y, 1 - \theta x - \theta y \rangle$ . Аналогично,

$$d\Phi = \begin{pmatrix} 0 & \gamma \\ \gamma & 0 \end{pmatrix} \oplus 0 \rightarrow d\Phi\mathbb{I} = \begin{pmatrix} 0 & \theta\gamma \\ \theta\gamma & 0 \end{pmatrix} \oplus 0 = \theta d\Phi.$$

Отсюда по предположенной монотонности  $Q_{\Phi\mathbb{I}}(d\Phi\mathbb{I}) \leq Q_{\Phi}(d\Phi)$  следует (8.1), что и требовалось.

Ввиду однородности (7.10) функцию  $c(x, y)$  естественно оценить на границе области задания, т.е. при  $x + y = 1$ ,  $y = 1 - x$ ,  $x, y > 0$ .

**ТЕОРЕМА 8.2.** *Если риманова метрика  $\rho$  монотонна в категории, то выполнены неравенства*

$$c\left(\frac{1}{2}, \frac{1}{2}\right) \leq c(x, 1 - x) \leq c\left(\frac{1}{2}, \frac{1}{2}\right) \frac{1}{4x(1 - x)}, \quad (8.3)$$

$$x \geq x' \geq \frac{1}{2} \implies c(x, 1 - x) \geq c(x', 1 - x'), \quad (8.4)$$

$$\frac{c(x, 1 - x)}{1 - x} \geq \frac{dc(x, 1 - x)}{dx}. \quad (8.5)$$

Доказательство всех этих неравенств проводится по одному плану, подбором соответствующего марковского эндоморфизма совокупности  $\mathfrak{S}(\mathfrak{B}(\mathcal{H}_2))$ . Пусть  $x > 1/2 > 1 - x$ . Рассмотрим эндоморфизм  $\mathbb{I}$  с четыре-матрицей второго порядка

$$\left( \begin{array}{c|c} \begin{array}{cc} 1 - \vartheta & 0 \\ 0 & \vartheta \end{array} & \begin{array}{cc} 0 & 1 - \vartheta \\ \vartheta & 0 \end{array} \\ \hline \begin{array}{cc} 0 & \vartheta \\ 1 - \vartheta & 0 \end{array} & \begin{array}{cc} \vartheta & 0 \\ 0 & 1 - \vartheta \end{array} \end{array} \right).$$

Она вполне положительна при всех  $\vartheta \in [0, 1]$ . На операторы с матрицей  $\text{diag} \langle x, 1 - x \rangle$  действует по формуле  $[\text{diag} \langle x, 1 - x \rangle] \Pi = \text{diag} \langle x - \vartheta \Delta, 1 - x + \vartheta \Delta \rangle$ , где  $\Delta = x - (1 - x) = 2x - 1 > 0$  по предположению. А для вещественного дифференциала

$$d\Phi = \begin{pmatrix} 0 & \gamma \\ \gamma & 0 \end{pmatrix}, \quad d\Phi\Pi = d\Phi.$$

Так как при изменении  $\vartheta$  от 0 до 1  $x' = x - \vartheta\Delta$  пробегает все значения в интервале  $[x, 1 - x]$ , то отсюда по монотонности  $Q_{\Phi\Pi}(d\Phi\Pi) \leq Q_{\Phi}(d\Phi)$  следует (8.4). В частности, при  $\vartheta = -1/2$  получаем левое неравенство (8.3).

Рассмотрим теперь при  $0 < x < 1$  эндоморфизм  $\mathcal{K}$  с матрицей

$$\left( \begin{array}{cc|cc} x & 0 & 0 & \sqrt{x(1-x)} \\ 0 & 1-x & \sqrt{x(1-x)} & 0 \\ \hline 0 & \sqrt{x(1-x)} & x & 0 \\ \sqrt{x(1-x)} & 0 & 0 & 1-x \end{array} \right)$$

и его действие на операторы и дифференциалы указанного выше вида. Имеем  $[\text{diag} \langle 1/2, 1/2 \rangle] \mathcal{K} = \text{diag} \langle x, 1 - x \rangle$ ,  $d\Phi\mathcal{K} = 2\sqrt{x(1-x)} d\Phi$ . Выписанная четырехматрица вполне положительна по критерию (3.4). Отсюда по монотонности метрики и свойству (7.10) получаем левое неравенство (8.3).

Наконец, рассмотрим эндоморфизм  $\mathcal{I}$  с вполне положительной при  $0 < \vartheta < 1$  матрицей

$$\left( \begin{array}{cc|cc} 1 & 0 & 0 & \sqrt{1-\vartheta} \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & \vartheta & 0 \\ \sqrt{1-\vartheta} & 0 & 0 & 1-\vartheta \end{array} \right).$$

На интересующие нас операторы и дифференциалы он действует по формулам

$$\begin{aligned} [\text{diag} \langle x, 1 - x \rangle] \mathcal{I} &= \text{diag} \langle x + \vartheta(1 - x), (1 - x)(1 - \vartheta) \rangle, \\ d\Phi\mathcal{I} &= \sqrt{1 - \vartheta} d\Phi. \end{aligned}$$

Отсюда по монотонности получаем

$$\begin{aligned} c(x, 1 - x) &\geq (1 - \vartheta) \cdot c(x + \vartheta(1 - x), 1 - x - \vartheta(1 - x)), \\ \vartheta c(x + \delta, 1 - x - \delta) &\geq c(x + \delta, 1 - x - \delta) - c(x, 1 - x), \end{aligned}$$

где  $\delta = \vartheta(1 - x) > 0$ . Поделив обе части на  $\delta$ , приходим к неравенству

$$\frac{c(x + \delta, 1 - x - \delta)}{1 - x} \geq \frac{c(x + \delta, 1 - x - \delta) - c(x, 1 - x)}{\delta}.$$

Выражение справа неотрицательно. Выражение слева ограничено по (8.3) при любом  $\delta < 1 - x$ , т.е.  $\vartheta < 1$ , и монотонно убывает по (8.4) при  $\delta \downarrow 0$ . Если предел справа существует, то он является правой производной и для нее выполнено неравенство (8.5). Так как функция  $c(x, 1 - x)$  монотонна, то она имеет

производную почти всюду. В исключительных точках выполняется аналогичное неравенство для локальной константы Липшица. Теорема доказана.

Так как совокупность  $\mathfrak{S}(\mathfrak{B})$  выпукла, то можно рассматривать значения квадратичных форм  $Q_\Phi(\cdot)$  не только от дифференциалов, но и от конечных приращений. Для классической фишеровской информации (т.е. для  $\Delta\Phi$ , коммутирующих с  $\Phi$ )

$$Q_{\Phi \otimes \Psi}(\Delta[\Phi \otimes \Psi]) = Q_\Phi(\Delta\Phi) + Q_\Psi(\Delta\Psi) + Q_\Phi(\Delta\Phi) \cdot Q_\Psi(\Delta\Psi). \quad (8.6)$$

**ТЕОРЕМА 8.3.** *Если задающая монотонную риманову метрику квадратичная форма  $Q$  удовлетворяет обобщенному принципу сложения информации (8.6) при любых самосопряженных приращениях  $\Delta\Phi$  и  $\Delta\Psi$ , то отвечающая ей функция  $c(x, y)$  удовлетворяет функциональному уравнению*

$$c(pu, qv) + c(pv, qu) = 2c(p, q) \cdot c(u, v) \quad (8.7)$$

при всех допустимых значениях аргументов

$$0 < p, q < p + q \leq 1, \quad 0 < u, v < u + v \leq 1.$$

**Доказательство.** Пусть  $\Phi = \text{diag} \langle p_1, p_2, \dots \rangle$ ,  $\Psi = \text{diag} \langle q_1, q_2, \dots \rangle$ , приращение  $\Delta\Phi$  сводится к недиагональному возмущению  $\Delta\varphi_j^i = \overline{\Delta\varphi_j^i}$ , аналогично  $\Delta\Psi$ :  $\Delta\psi_l^k = \overline{\Delta\psi_l^k}$ . Поскольку

$$\Delta(\Phi \otimes \Psi) = \Phi \otimes (\Delta\Psi) + (\Delta\Phi) \otimes \Psi + (\Delta\Phi) \otimes (\Delta\Psi),$$

$$Q_{\Phi \otimes \Psi}(\Delta[\Phi \otimes \Psi]) = 2 \sum_{\alpha} c(p_{\alpha}q_k, p_{\alpha}q_l) |p_{\alpha}\Delta\psi_l^k|^2 + 2 \sum_{\beta} c(p_iq_{\beta}, p_jq_{\beta}) |q_{\beta}\Delta\varphi_j^i|^2 + \\ + 2 [c(p_lq_k, p_jq_l) + c(p_iq_l, p_jq_k)] \cdot |\Delta\varphi_j^i|^2 \cdot |\Delta\psi_l^k|^2,$$

так как  $\sum p_{\alpha} = \sum q_{\beta} = 1$ , то, вынося  $p_{\alpha}$  и  $q_{\beta}$  по правилу (7.10), мы приводим первые два правых слагаемых к виду

$$2c(q_k, q_l) |\Delta\psi_l^k|^2 = Q_{\Psi}(\Delta\Psi), \quad 2c(p_i, p_j) |\Delta\varphi_j^i|^2 = Q_{\Phi}(\Delta\Phi).$$

Отсюда, если (8.6) выполняется, последняя сумма в квадратных скобках должна приводиться к произведению, т.е. соотношение (8.7) должно выполняться. Нетрудно выкладкой проверить, что условие (8.7) является и достаточным для выполнения (8.6).

**ЛЕММА 8.4.** *Функциональному уравнению (8.7) удовлетворяют функции*

$$c(u, v) = c[u^{\alpha}v^{-\alpha} + u^{-\alpha}v^{\alpha}] (uv)^{\beta} \quad (8.8)$$

при любых вещественных  $\alpha$  и  $\beta$ . Они вещественно однородны в минус первой степени при  $\beta = -1/2$ . При  $0 \leq \alpha \leq 1/2$  эти последние удовлетворяют неравенству (8.3).

Доказательство следует из тождества

$$\begin{aligned} [u^\alpha v^{-\alpha} + u^{-\alpha} v^\alpha] [p^\alpha q^{-\alpha} + p^{-\alpha} q^\alpha] = \\ = [(pu)^\alpha (qv)^{-\alpha} + (pu)^{-\alpha} (qv)^\alpha] + [(pv)^\alpha (qu)^{-\alpha} + (pv)^{-\alpha} (qu)^\alpha]. \end{aligned}$$

В частных случаях имеем (см. [9])

$$\begin{aligned} Q_\Phi(d\Phi) &= c \operatorname{tr}[\Phi^{-1}(d\Phi)^2], & \alpha &= \frac{1}{2}, \\ Q_\Phi(d\Phi) &= c \operatorname{tr}[(d\Phi)\Phi^{-1/2}(d\Phi)\Phi^{-1/2}], & \alpha &= 0. \end{aligned}$$

**9.** Класс функций  $c(p, q)$ , заданных в замкнутой области

$$0 \leq p, q \leq p + q \leq 1 \quad (9.1)$$

и удовлетворяющих необходимым условиям

$$c(p, p) = cp^{-1}, \quad (9.2)$$

$$c(p, q) = c(q, p) \quad (9.3)$$

(полученным в теореме 7.2) и условию (7.10), не исчерпывается решениями функционального уравнения (8.7).

**ТЕОРЕМА 9.1.** *Следующие функции  $c(p, q)$  удовлетворяют условиям (9.1), (9.2), (9.3), (7.10) и задают некоторые римановы метрики:*

$$\frac{2}{p+q} : ds^2 = \operatorname{tr}[\Phi(dR)^2], \quad d\Phi = (dR)\Phi + \Phi(dR); \quad (9.4)$$

$$\left(\frac{2}{\sqrt{p} + \sqrt{q}}\right)^2 : ds^2 = 4 \operatorname{tr}(d\sqrt{\Phi})^2, \quad s = 2 \arccos \operatorname{tr}(\sqrt{\Phi}\sqrt{\Psi}); \quad (9.5)$$

$$\frac{\ln p - \ln q}{p - q} : ds^2 = \operatorname{tr}[(d \ln \Phi)(d\Phi)]; \quad (9.6)$$

$$\left(\frac{\ln p - \ln q}{p - q}\right)^2 \frac{p+q}{2} : ds^2 = \operatorname{tr}[\Phi(d \ln \Phi)^2]; \quad (9.7)$$

$$\frac{1}{\sqrt{pq}} : ds^2 = \operatorname{tr}[\Phi^{-1/2}(d\Phi)\Phi^{-1/2}(d\Phi)]; \quad (9.8)$$

$$\frac{p^{2\alpha} + q^{2\alpha}}{2(pq)^{\alpha+1/2}} : ds^2 = \operatorname{tr}[\Phi^{-\alpha-1/2}(d\Phi)\Phi^{\alpha-1/2}(d\Phi)]; \quad (9.9)$$

$$\frac{1}{2} \left(\frac{1}{p} + \frac{1}{q}\right) : ds^2 = \operatorname{tr}[\Phi^{-1}(d\Phi)^2]. \quad (9.10)$$

Все эти функции задают некоторые средние обратных величин  $p$  и  $q$  и выписаны в порядке возрастания: от обратного к среднему арифметическому  $p$  и  $q$  до обратного к их среднему гармоническому (т.е. среднему арифметическому их обратных величин). При этом среднее (9.9) растет с увеличением  $\alpha$ .

**Доказательство.** Проверка условий (9.2), (9.3) и (7.10) элементарна. Разве что для (9.6) и (9.7) надо раскрыть неопределенности. Мы опускаем также проверку монотонности выписанной таблицы. В самом сложном технически сравнении (9.7) и (9.8) надо перейти от  $p$  и  $q$  к их логарифмам и сравнить возникающие после преобразования ряды с положительными слагаемыми.

**ЛЕММА 9.2.** *Функции  $c(p, q)$ , описанные формулами (9.4)–(9.10), удовлетворяют неравенствам (8.3)–(8.5) теоремы 8.2.*

**Доказательство.** Выполнение (8.3) вытекает из последнего утверждения теоремы 9.1. Выполнение (8.4) вытекает из легко проверяемой монотонности  $c(x, 1 - x)$  при  $x \geq 1/2$ , сводящейся к проверке  $dc(x, 1 - x)/dx \geq 0$  в области  $1/2 < x < 1$ . Наконец, неравенство (8.5) проверяется самостоятельно.

**10.** Для проверки инвариантности построенных римановых метрик мы используем описание семейств состояний, конгруэнтных в марковской категории; см. [25, 1].

**ЛЕММА 10.1.** *Если два семейства состояний  $\{\Phi_\theta^{(1)}\} \subseteq \mathfrak{S}_1$  и  $\{\Phi_\theta^{(2)}\} \subseteq \mathfrak{S}_2$ ,  $\theta \in \Theta$ , конгруэнтны относительно узкой марковской категории,*

$$\forall \theta \in \Theta \quad \Phi_\theta^{(1)} \Pi^{12} = \Phi_\theta^{(2)}, \quad \Phi_\theta^{(2)} \Pi^{21} = \Phi_\theta^{(1)}, \quad (10.1)$$

то  $\{\Phi_\theta^{(1)}\}$  и  $\{\Phi_\theta^{(2)}\}$  суть семейства состояний, соответственно стационарных относительно марковских преобразований  $\Pi^{12}\Pi^{21}$  и  $\Pi^{21}\Pi^{12}$ . Конгруэнтность  $(\Pi^{12}, \Pi^{21})$  продолжается на совокупности  $\mathfrak{S}_1^{(\Pi)} \subseteq \mathfrak{S}_1$  и  $\mathfrak{S}_2^{(\Pi)} \subseteq \mathfrak{S}_2$  всех стационарных состояний этих марковских преобразований.

**Доказательство.** Нуждается в проверке только последнее утверждение. Пусть

$$\Psi^{(1)} \Pi^{12} \Pi^{21} = \Psi^{(1)} \in \mathfrak{S}_1. \quad (10.2)$$

Обозначим  $\Psi^{(2)} = \Psi^{(1)} \Pi^{12}$ . Тогда

$$\Psi^{(2)} (\Pi^{21} \Pi^{12}) = \Psi^{(1)} (\Pi^{12} \Pi^{21}) \Pi^{12} = \Psi^{(2)} \in \mathfrak{S}_2.$$

Строение семейства  $\mathfrak{S}^{(\Pi)}$  всех состояний, стационарных относительно марковского оператора  $\Pi$ , изучалось нами в [1] и [25] для конечномерных гильбертовых пространств  $\mathcal{H}$  и в [2] и [16] для сепарабельных  $\mathcal{H}$ . Из проведенных там рассуждений для  $\dim \mathcal{H} < \infty$  вытекает следующая

**ТЕОРЕМА 10.2.** *С каждым марковским оператором  $\Pi$  на  $\mathfrak{S}(\mathcal{H})$  связывается разложение*

$$\mathcal{H} = \mathcal{N}_0 \oplus \bigoplus_{u=1}^w \mathcal{N}_u, \quad \dim \mathcal{N}_0 \geq 0, \quad \dim \mathcal{N}_u \geq 1, \quad w \geq 1, \quad (10.3)$$

так что в соответствующем базисе матрица каждого  $\Pi$ -стационарного состояния  $\Phi$  имеет вид

$$\Phi = 0_0 \oplus \bigoplus_{u=1}^w q_u \Phi_u; \quad \sum_{u=1}^w q_u = 1 \quad \forall u, \quad q_u \geq 0. \quad (10.4)$$

При этом если отвечающий  $\Phi$  набор  $\{\Phi_1, \dots, \Phi_w\}$  фиксировать и варьировать веса  $q_u$ , то получающиеся матрицы также будут отвечать  $\Pi$ -стационарным состояниям, так что

$$\mathfrak{S}^\Pi = \text{Join} \{ \mathfrak{S}^{\Pi_1}, \dots, \mathfrak{S}^{\Pi_w} \}. \quad (10.5)$$

С каждым сектором  $\mathcal{N}$  связывается своя числовая система  $\mathbb{K} = \mathbb{R}, \mathbb{C}, \mathbb{H}$  или  $\mathbb{K} = \text{Cl}_\mu$ , а также формальная размерность  $t$  над системой  $\mathbb{K}$ . В случае чисел Клиффорда  $\text{Cl}_\mu$  обязательно  $t = 2$ ,  $\mathbb{K}$  и  $t$  определяются совокупностью атомарных стационарных мер сектора  $\mathcal{N}$  и ее симметриями.

При  $\mathbb{K} = \mathbb{R}, \mathbb{H}, \text{Cl}_{2\nu}$  существует разложение

$$\mathcal{N} = \mathbb{C}^m \otimes \mathbb{C}^{2\nu} \otimes \mathbb{C}^t \quad (10.6)$$

в тензорное произведение подпространств, в соответствующих базисах которых матрица состояния  $\Phi_u$  из сектора  $\mathcal{N}_u$  имеет вид

$$\Phi = \left( \gamma_0^{ij} \mathcal{B}_{\nu,0} + \sum_{s=1}^{2\nu} \gamma_s^{ij} \mathcal{B}_{\nu,s} \right)_{i,j=1}^m \otimes \Psi, \quad (10.7)$$

где  $\Psi = \Psi_u$  — фиксированная для данного  $\Pi$  неотрицательная эрмитова матрица со следом 1, т.е. фиксированное состояние на  $\mathfrak{B}(\mathcal{H}_t)$ ,

$$\Gamma = (z^{ij}) = \left( \gamma_0^{ij} + \sum_{s=1}^{2\nu} e_s \gamma_s^{ij} \right)_{i,j=1}^m \quad (10.8)$$

— любая нормированная (т.е.  $\text{tr} \Gamma = 1$ ) неотрицательная эрмитова матрица над числовой системой  $\mathbb{K}_u$  с единицами  $1, e_1, \dots, e_{2\nu}$ ,  $\mathcal{B}_{\nu,s}$  — канонические представления единиц числовой системы  $\mathbb{K}_u$ .

При  $\mathbb{K} = \mathbb{C}, \text{Cl}_{2\nu+1}$  существует более сложное разложение каждого сектора

$$\mathcal{N} = \mathbb{C}^m \otimes \mathbb{C}^{2\nu} \otimes (\mathbb{C}^{t'} \oplus \mathbb{C}^{t''}), \quad (10.9)$$

в фиксированных соответствующих базисах которого матрицы  $\Phi_u$  из сектора  $\mathcal{N}_u$  взаимно однозначно представляются в виде

$$\Phi = \left( \gamma_0^{ij} \mathcal{B}_{\nu,0} + \sum_{s=1}^{2\nu+1} \gamma_s^{ij} \mathcal{B}_{\nu,s} \right) \otimes \Psi' \oplus \left( \gamma_0^{ij} \mathcal{B}_{\nu,0} - \sum_{s=1}^{2\nu+1} \gamma_s^{ij} \mathcal{B}_{\nu,s} \right) \otimes \Psi'', \quad (10.10)$$

с фиксированными эрмитовыми  $\Psi'$  и  $\Psi''$ ,

$$\Psi', \Psi'' \geq 0, \quad \text{tr} \Psi' + \text{tr} \Psi'' = 1, \quad (10.11)$$

соответственно двум каноническим представлениям числовой системы  $\mathbb{K}_u$  с единицами  $1, e_1, \dots, e_{2\nu+1}$ .

Доказательство сформулированного утверждения дано в § 5 нашего препринта [1]. Оно использует тот факт, что  $\Pi$ -стационарные заряды  $\Xi$ ,

$$\Xi \Pi = \Xi, \quad (10.12)$$

образуют расщеп, т.е. множество  $\Xi$ , удовлетворяющих (10.12), образует линейал, в котором из (10.12) вытекает

$$\Xi^+ \Pi = \Xi^+, \quad \Xi^- \Pi = \Xi^-, \tag{10.13}$$

где  $\Xi^+$  и  $\Xi^-$  — положительная и отрицательная части заряда  $\Xi$ ; см. подробнее [1, 2].

**ЛЕММА 10.3.** Семейство  $\mathfrak{S}^\Pi$  конгруэнтно в узкой марковской категории семейства  $\mathfrak{S}_{\text{red}}^\Pi = \text{Join} \{ \mathfrak{S}_{\text{red}}^{\Pi_1}, \dots, \mathfrak{S}_{\text{red}}^{\Pi_w} \}$ , получаемому отбрасыванием множителей  $\mathbb{C}^t$  в (10.6) и (10.9) и множителей  $\Psi$  в (10.7) и (10.10).

Доказательство достаточно провести для каждого сектора отдельно. Как известно, свертка по части индексов и тензорное умножение на матрицу плотности задают марковские отображения. В секторе  $\mathcal{M}_0$  достаточно задать отображение на редукцию правилом  $\Phi \rightarrow \Phi_1^{(0)} \rightarrow \Phi_{1\text{red}}^{(0)}$ , где  $\Phi_1^{(0)} \in \mathfrak{S}^{\Pi_1}$  фиксировано.

**ОПРЕДЕЛЕНИЕ.** Семейство  $\mathfrak{S}_{\text{red}}^\Pi$  назовем *минимальным представлением семейства*  $\mathfrak{S}^\Pi$ .

**ТЕОРЕМА 10.4.** Если дифференциальная квадратичная форма  $Q_\Phi(d\Phi)$  вида (7.7) задается функцией  $c(p, q)$ , удовлетворяющей условиям (9.1)–(9.3), то она инвариантна в широкой марковской категории.

**Доказательство.** Проверим, что квадрат дифференциала дуги сохраняется при редукции на минимальное представление. Будем считать, что все матрицы  $\Psi_u$  приведены к диагональному виду, все матрицы  $\Gamma_u = \Gamma(\Phi_u)$  приведены к диагональному (и, следовательно, вещественному) виду. После перехода от  $\mathbb{K}$ -чисел к их каноническим  $\mathbb{C}$ -представлениям (4.7) и (4.8) из [1] будем иметь

$$\Gamma(\Phi_u) = \bigoplus_{j=1}^m \gamma_j \mathbf{I}_{2^{2\nu(u)}}, \quad \Gamma(d\Phi_u) = (h_u^{kl})_{k,l=1}^{m2^{2\nu(u)}}. \tag{10.14}$$

Тогда для  $\Phi = \bigoplus_u \Gamma(\Phi_u) \otimes \Psi_u$ ,  $d\Phi = \bigoplus_u \Gamma(d\Phi_u) \otimes \Psi_u$ ,  $d\Psi_u = 0$  имеем

$$\begin{aligned} Q_\Phi(d\Phi) &= \sum_u \sum_{k,l=1}^{m2^{2\nu}} \sum_j c(p_{uk} \psi_j, p_{ul} \psi_j) |h_u^{kl} \psi_j|^2 = \\ &= \sum_u \sum_{k,l} c(p_{uk}, p_{ul}) |h_u^{kl}|^2 \text{tr } \Psi_u = Q_{\Phi_{\text{red}}}(d\Phi_{\text{red}}), \end{aligned}$$

где каждое значение  $u$  при нечетном  $\gamma$  может быть представлено не одним, а двумя слагаемыми с  $h_u^{kl'} = \pm h_u^{kl''}$ , и  $\psi_j = \psi_u^{jj}$ . Так как конгруэнтные семейства одновременно приводятся к каноническому представлению, поскольку оно задается соответствующим набором атомарных состояний, свойства которого инвариантны относительно конгруэнтности, то их минимальные представления могут быть выбраны одинаковыми. Остается доказать, что форма  $Q_\Phi(d\Phi)$  не меняется при унитарном преобразовании пространства, сохраняющем диагональный вид матрицы  $\Phi$ . Очевидно, достаточно рассмотреть движения, затрагивающие только

одно из собственных подпространств  $\Phi$ , например первое. Имеем

$$\begin{aligned} \sum_{k,l=1}^n c(p_k, p_l) |d\varphi_l^k|^2 &= c(p, p) \sum_{k,l=1}^m |d\varphi_l^k|^2 + \sum_{l=m+1}^n c(p, p_l) \sum_{k=1}^m |d\varphi_l^k|^2 + \\ &+ \sum_{k=m+1}^n c(p_k, p) \sum_{l=1}^m |d\varphi_l^k|^2 + \sum_{k,l>m}^n c(p_k, p_l) |d\varphi_l^k|^2. \end{aligned}$$

Все слагаемые, очевидно, сохраняются.

**11.** Каждое распределение вероятностей  $P$  на алгебре  $\mathcal{A}_n$  всех подмножеств множества  $\Omega_n \{\omega_1, \dots, \omega_n\}$  всех исходов  $\omega$  задается вектором-строкой  $\mathbf{p} = \langle\langle p_1, \dots, p_n \rangle\rangle$  вероятностей атомов,  $\forall j \ p_j = P\{\omega_j\}$ . Для распределения вероятностей

$$\forall j \ p_j \geq 0; \quad \sum_{j=1}^n p_j = 1. \quad (11.1)$$

Обозначим через  $\mathbf{e}(k)$ ,  $k = 1, \dots, n$ , вектор распределения, сосредоточенного на  $\omega_k$ :

$$e_j(k) = \delta_j^{(k)}. \quad (11.2)$$

Тогда в каждой внутренней точке  $\mathbf{p}$  симплекса (11.1) возникает переполненный (на один вектор) базис

$$\mathbf{e}(k) - \mathbf{p}, \quad k = 1, \dots, n, \quad (11.3)$$

и связанные с ним реперы

$$p_k \mathbf{e}(k) - p_k \mathbf{p}, \quad k = 1, \dots, n, \quad (11.4)$$

$$(1 - p_k)^{-1} [\mathbf{e}(k) - \mathbf{p}] = \mathbf{e}(k) - \mathbf{q}(k), \quad (11.5)$$

где вектор  $\mathbf{q}(k)$  — вектор вероятностей атомов для условного распределения  $\mathbf{P}\{\cdot | \Omega - \omega_k\}$ , лежащий в  $k$ -й грани симплекса. Для каждого фиксированного вектора  $\mathbf{q}(k)$  чевиана симплекса

$$\mathbf{p}(t) = t\mathbf{e}(k) + (1 - t)\mathbf{q}(k) = \mathbf{q}(k) + t[\mathbf{e}(k) - \mathbf{q}(k)] \quad (11.6)$$

при  $0 \leq t \leq 1$  задает семейство распределений с одним и тем же вектором условных вероятностей  $\mathbf{q}(k)$ . Расслоение симплекса на чевианы, исходящие из  $\mathbf{e}(k)$ , задает внутри симплекса поле касательных векторов — дифференцирование по натуральному параметру  $t$ :

$$(Z_k f)_{\mathbf{p}} = \left. \frac{d}{dt} f(\mathbf{p}(t)) \right|_{\mathbf{p}(t)=\mathbf{p}}. \quad (11.7)$$

Будем считать, что гладкая функция  $f(\cdot)$  задана на всем положительном ортанте

$$\mathbf{z} : \forall j \ z_j \geq 0, \quad (11.8)$$

а не только на симплексе (11.1); например, продолжена с (11.1) на (11.8) по однородности:

$$f(hp_1, \dots, hp_n) = h \cdot f(p_1, \dots, p_n) \quad \forall h > 0. \quad (11.9)$$

Тогда для дифференцирования вдоль любой гладкой кривой  $\mathbf{p}(\theta)$  справедливо

$$\frac{d}{d\theta} f(\mathbf{p}(\theta)) = \frac{\partial f}{\partial p_1} \frac{\partial p_1}{\partial \theta} + \dots + \frac{\partial f}{\partial p_n} \frac{\partial p_n}{\partial \theta}. \quad (11.10)$$

Эту связь мы будем символически записывать

$$\frac{d}{d\theta} \leftrightarrow \langle\langle p'_1(\theta), \dots, p'_n(\theta) \rangle\rangle = \mathbf{p}'(\theta). \quad (11.11)$$

Тогда для дифференцирования  $Z_k$  по натуральному параметру  $t$  вдоль чевиан  $k$ -го семейства

$$Z \leftrightarrow \mathbf{e}(k) - \mathbf{q}(k) = (1 - p_k)^{-1} [\mathbf{e}(k) - \mathbf{p}]. \quad (11.12)$$

Перепараметризуем чевиану так, чтобы

$$\frac{d}{ds} = t(1-t) \frac{d}{dt}, \quad ds = \frac{dt}{t} + \frac{dt}{1-t}. \quad (11.13)$$

Векторные поля  $X_k$  дифференцирований по параметру  $s$ ,

$$X_k = \frac{d}{ds_k} \leftrightarrow p_k \mathbf{e}(k) - p_k \mathbf{p}, \quad (11.14)$$

оказываются более удобными для описания линейных связностей. В силу связи (11.14)

$$0 = \mathbf{p} - \mathbf{p} = \sum_{k=1}^n [p_k \mathbf{e}(k) - p_k \mathbf{p}] \leftrightarrow \sum_{k=1}^n X_k,$$

откуда

$$\sum_{k=1}^n X_k = 0. \quad (11.15)$$

ЛЕММА 11.1. *Любой касательный вектор  $V = \frac{d}{d\theta}$  такой, что*

$$(Vf)_p = \sum_{k=1}^n v_k(\mathbf{p}) \left. \frac{\partial f}{\partial p_k} \right|_{\mathbf{p}}, \quad \sum_{k=1}^n v_k = 0, \quad (11.16)$$

удобно раскладывать по простейшему переполненному базису  $(Y_j)_{\mathbf{p}} \leftrightarrow \mathbf{e}(j) - \mathbf{p}$ :

$$V = \sum_{j=1}^n v_j Y_j. \quad (11.17)$$

Доказательство. В самом деле,

$$\sum_{j=1}^n v_j Y_j \leftrightarrow \sum_{j=1}^n v_j [\mathbf{e}(j) - \mathbf{p}] = \sum_{j=1}^n v_j \mathbf{e}(j) - \mathbf{p} \sum_{j=1}^n v_j = \sum_{j=1}^n v_j \mathbf{e}(j) = \langle v_1, \dots, v_n \rangle \leftrightarrow V.$$

В остальных реперах одно из представлений можно получить подстановкой. Например,

$$V = \sum_{j=1}^n \frac{v_j}{p_j} X_j = \sum_{j=1}^n \xi_j^V(\mathbf{p}) X_j, \quad \sum_{j=1}^n \xi_j^V(\mathbf{p}) p_j = 0. \quad (11.18)$$

Ввиду (11.15) общий вид разложения —

$$V = \sum_{j=1}^n \left( \frac{v_j}{p_j} - c \right) X_j \quad \forall c \in \mathbb{R}. \quad (11.19)$$

Построим теперь двумерную поверхность, расслаивающуюся одновременно на чевианы  $j$ -го и  $k$ -го семейств. Для упрощения обозначений возьмем  $j = 1$ ,  $k = 2$ .

ЛЕММА 11.2. Семейство  $P_s(\cdot)$ ,  $\mathbf{s} = \langle s_1, s_2 \rangle$ , описываемое семейством векторов  $\mathbf{p}(s_1, s_2) \in \text{Int Cap}(\Omega_n, \mathcal{A}_n)$

$$\begin{aligned} p_1(s_1, s_2) &= p_1(0, 0) \exp[s_1 - \Psi(s_1, s_2)], \\ p_2(s_1, s_2) &= p_2(0, 0) \exp[s_2 - \Psi(s_1, s_2)], \\ p_i(s_1, s_2) &= p_i(0, 0) \exp[-\Psi(s_1, s_2)], \quad i \geq 3, \\ \Psi(s_1, s_2) &= \ln \left[ p_1(0, 0) \exp s_1 + p_2(0, 0) \exp s_2 + \sum_{i \geq 3} p_i(0, 0) \right], \end{aligned} \quad (11.20)$$

при фиксированном значении  $s_2 = \sigma$  задает чевиану первого семейства в канонической параметризации, связанной с натуральной координатой  $t_1$  соотношением (11.13). Аналогичная конструкция задает чевианы второго семейства.

СЛЕДСТВИЕ. Коммутатор базисных полей

$$[X_j, X_k] = X_j X_k - X_k X_j = 0 \quad \forall j, k. \quad (11.21)$$

Доказательство. Обозначим  $t = t(s_1, \sigma) = p_1(s_1, \sigma)$  и введем вектор  $\mathbf{q}_{(1)} = [1 - t(0, \sigma)]^{-1} \{ \mathbf{p}(0, \sigma) - t(0, \sigma) \mathbf{e}(1) \}$ ,  $q_1(1) = 0$ . Легко проверить покомпонентно, что

$$\forall s_1 \quad [1 - t(s_1, \sigma)]^{-1} \{ \mathbf{p}(s_1, \sigma) - t(s_1, \sigma) \mathbf{e}(1) \} = \mathbf{q}(1). \quad (11.22)$$

Таким образом, обратно,  $\mathbf{p}(s_1, \sigma)$  образуют чевиану

$$\mathbf{p}(s_1, \sigma) = t(s_1, \sigma) \mathbf{e}(1) + [1 - t(s_1, \sigma)] \mathbf{q}(1). \quad (11.23)$$

Найдем связь между дифференцированием по  $s_1$  и по  $t$ . Так как

$$\frac{d\Psi(s_1, \sigma)}{ds_1} = p_1, \quad (11.24)$$

то, дифференцируя покомпонентно, находим  $\mathbf{p}'_{s_1} = p_1 \mathbf{e}(1) - p_1 \mathbf{p}$ , откуда  $\frac{\partial}{\partial s_1} = X_1$  в силу (11.13) и (11.14). Аналогично,  $\frac{\partial}{\partial s_2} = X_2$ . На поверхности (11.20)  $\frac{\partial^2 \mathbf{p}}{\partial s_1 \partial s_2} = \frac{\partial^2 \mathbf{p}}{\partial s_2 \partial s_1}$ , т.е.  $X_1 X_2 = X_2 X_1$ , что и требовалось доказать.

**12.** Любая пара взаимно дизъюнктивных распределений вероятностей конгруэнтна любой другой такой паре (см. следствие из теоремы 6.2 в [13]). Поэтому любая чевиана (11.6) конгруэнтна простейшей совокупности  $\text{Car}_2$ :

$$\{\langle p_1, p_2 \rangle : p_1 = t, p_2 = 1 - t, 0 \leq t \leq 1\} = \text{Car}_2, \tag{12.1}$$

— отрезку с концами  $\mathbf{e}(1)$  и  $\mathbf{e}(2)$ . Этот отрезок обязан быть геодезической в любой инвариантной линейной связности. Имеем по (11.13) и (11.14)

$$n = 2, \quad X_1 = -X_2 = X = t(1 - t) \frac{d}{dt}. \tag{12.2}$$

Согласно теореме 12.2 из [13] для связности  ${}^\gamma \nabla$  выполнены соотношения

$${}^\gamma \nabla x_k X_j = -\gamma [p_j X_k + p_k X_j], \quad j \neq k, \tag{12.3}$$

$${}^\gamma \nabla x_j X_j = \gamma(1 - 2p_j) X_j. \tag{12.4}$$

Для геодезической (12.1) оба эти соотношения ввиду (12.2) эквивалентны правилу переноса

$${}^\gamma \nabla_X X = \gamma(1 - 2t) X. \tag{12.5}$$

**ТЕОРЕМА 12.1.** Поле параллельных вдоль геодезической (12.1) касательных векторов  ${}^\gamma V$  в инвариантной линейной связности  ${}^\gamma \nabla$  задает поле дифференцирований

$$\frac{d}{d\theta} = V = [t(1 - t)]^{-\gamma} X = [t(1 - t)]^{-\gamma+1} \frac{d}{dt}. \tag{12.6}$$

*Доказательство.* Вычислим  ${}^\gamma \nabla_V V$ , используя (12.5), (12.2) и основные свойства символа  $\nabla$  (см. [23, гл. 1, § 4] или [13, § 3]):

$$\begin{aligned} \nabla_V V &= \nabla_{[\ ]^{-\gamma} X} [\ ]^{-\gamma} X = [\ ]^{-\gamma} \nabla_X [\ ]^{-\gamma} X = [\ ]^{-\gamma} (X [\ ]^{-\gamma}) X + [\ ]^{-2\gamma} \nabla_X X = \\ &= -[t(1 - t)]^{-\gamma} t(1 - t) \gamma [t(1 - t)]^{-\gamma-1} (1 - 2t) X + [t(1 - t)]^{-2\gamma} \gamma(1 - 2t) X = 0. \end{aligned}$$

Из доказанной теоремы нетрудно получить геодезические параметризации при разных  $\gamma$ . При  $\gamma = 1$   ${}^1 V = \frac{d}{dt}$ ,  $d\theta^1 = dt$ ,  $\theta^1 = t - t_0$ . При  $\gamma = 0$   ${}^0 V = X$ ,  $\theta^0 = s$ ,

$$ds = \frac{dt}{t} + \frac{dt}{1 - t}, \quad s = \ln t - \ln(1 - t), \quad t = \frac{\exp(s/2)}{\exp(s/2) + \exp(-s/2)}. \tag{12.7}$$

**СЛЕДСТВИЕ 1.** При  $\gamma = 0$  геодезический параметр  $s$  является каноническим параметром экспонентного семейства  $\{\langle p_1(s), p_2(s) \rangle\}$ ,

$$p_k(s) = \frac{\exp[(-1)^{k-1} s/2]}{\exp(s/2) + \exp(-s/2)}, \quad -\infty \leq s \leq \infty. \tag{12.8}$$

А его натуральным экспонентным параметром является  $t$ , т.е. геодезический параметр, отвечающий  $\gamma = 1$ .

При  $\gamma = 1/2$  линейная связность отвечает римановой метрике; см. [13, § 12].

СЛЕДСТВИЕ 2. При  $\gamma = 1/2$  геодезическим параметром является риманова длина дуги  $\sigma$ , отвечающая фишеровой квадратичной форме:

$$(d\sigma)^2 = \frac{(dt)^2}{t} + \frac{[d(1-t)]^2}{1-t} = 4(d\sqrt{t})^2 + 4(d\sqrt{1-t})^2, \quad (12.9)$$

$$d\sigma = [t(1-t)]^{-1/2} dt; \quad (12.10)$$

с точностью до выбора начала отсчета

$$\sigma - \sigma_0 = \arccos(1-2t) = 2 \arcsin \sqrt{t} = 2 \arccos \sqrt{1-t}. \quad (12.11)$$

СЛЕДСТВИЕ 3. Для  $\gamma$  общего вида геодезический параметр  $\theta$  выражается через  $t$  и  $\sigma$  квадратурами

$$d\theta = [t(1-t)]^{\gamma-1} dt = 2^{-2(\gamma-1)} \left[ \sin \frac{\sigma}{2} \right]^\gamma d\sigma = 2^{-2\gamma+1} [\sin \varphi]^\gamma d\varphi, \quad (12.12)$$

где  $\varphi = \sigma/2$ ,  $\sin \varphi = \sqrt{t}$ ,  $\cos \varphi = \sqrt{1-t}$ ,  $0 \leq \varphi \leq \pi/2$ .

Как отмечалось в начале этого параграфа, каждая чевиана конгруэнтна совокупности  $\text{Car}_2$ . Следовательно, она является геодезической линией при любом  $\gamma$ , и все соотношения (12.7)–(12.12) между геодезическими параметрами сохраняют силу и для нее.

**13.** Если инвариантная линейная связность  ${}^\gamma \nabla$  наследственна в тензорных степенях, то семейство тензорных квадратов распределений вероятностей совокупности  $\text{Car}_2$  должно быть геодезической линией (с тем же геодезическим параметром) в этой связности  ${}^\gamma \nabla$ , поскольку геодезической линией является сама  $\text{Car}_2$ . Выясним сейчас, при каких  $\gamma$  будет так.

Пусть даны два экземпляра множества  $\Omega_2 = \{\omega_1, \omega_2\}$ :  $\Omega'_2$  и  $\Omega''_2$ ;

$$\Omega'_2 \times \Omega''_2 = \{\omega'_1 \omega''_1, \omega'_1 \omega''_2, \omega'_2 \omega''_1, \omega'_2 \omega''_2\}.$$

Лексикографически перенумеровав исходы единой нумерацией, получаем  $\Omega'_2 \times \Omega''_2 = \Omega_4 = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ . Если  $\mathbf{p} = \langle\langle p_1, p_2 \rangle\rangle$  — распределение на  $(\Omega'_2, \mathcal{A}'_2)$ ,  $\mathbf{q} = \langle\langle q_1, q_2 \rangle\rangle$  — на  $(\Omega''_2, \mathcal{A}''_2)$ , то их произведение на  $(\Omega_4, \mathcal{A}_4)$  задается вектором

$$\mathbf{p} \otimes \mathbf{q} = \langle\langle p_1 q_1, p_1 q_2, p_2 q_1, p_2 q_2 \rangle\rangle. \quad (13.1)$$

Отсюда, полагая  $p_i = q_i$ ,  $i = 1, 2$ , квадрат совокупности  $\text{Car}_2$ , взятой в натуральной параметризации, будет описываться кривой

$$\mathbf{r}(t) = \langle\langle t^2, t(1-t), t(1-t), (1-t)^2 \rangle\rangle, \quad 0 \leq t \leq 1. \quad (13.2)$$

Поэтому семейство касательных векторов относительно исходной натуральной ( $\gamma = 1$ ) параметризации —

$$\dot{\mathbf{r}}(t) = \langle\langle 2t, 1-2t, 1-2t, -2+2t \rangle\rangle. \quad (13.3)$$

Относительно любой другой исходной  $\gamma$ -геодезической параметризации в силу связи (12.12)

$$\mathbf{r}'_{\theta}(t(\theta)) = \dot{\mathbf{r}}(t) \cdot t'_{\theta} = [t(1-t)]^{-\gamma+1} \dot{\mathbf{r}}(t). \quad (13.4)$$

На кривой  $\mathbf{r}(t(\theta))$  дифференцирование по параметру  $d/d\theta$  задает семейство касательных векторов, т.е.

$$(W)_{\mathbf{r}} \leftrightarrow [t(1-t)]^{-\gamma+1} \dot{\mathbf{r}}(t),$$

$$\begin{aligned} \langle\langle \omega_1, \omega_2, \omega_3, \omega_4 \rangle\rangle &= [t(1-t)]^{-\gamma+1} \dot{\mathbf{r}}(t) = \\ &= [r_2 r_3]^{-(\gamma-1)/2} \langle\langle 2\sqrt{r_1}, r_4 - r_1, r_4 - r_1, -2\sqrt{r_4} \rangle\rangle. \end{aligned} \quad (13.5)$$

Согласно (11.18) это продолжимое по (13.5) семейство касательных векторов  $W$  раскладывается по базисным полям  $X_1, X_2, X_3, X_4$  с коэффициентами, принимающими на кривой  $\mathbf{r}(t)$  вид

$$(W)_{\mathbf{r}(t)} = [t(1-t)]^{-\gamma} \{ (2-2t)X_1 + (1-2t)(X_2 + X_3) - 2tX_4 \}_{\mathbf{r}(t)}. \quad (13.6)$$

Это представление можно упростить по (11.19), взяв  $c(t) = 1 - 2t = r_4 - r_1$ :

$$\left. \frac{d}{d\theta} \right|_{\mathbf{r}(t)} = (W)_{\mathbf{r}(t)} = [t(1-t)]^{-\gamma} \{ X_1 - X_4 \}_{\mathbf{r}(t)}. \quad (13.7)$$

**ТЕОРЕМА 13.1.** Пусть ограничение векторного поля  $W$ , заданного по (13.5), на кривую  $\mathbf{r}(t) = \mathbf{p}(t) \otimes \mathbf{p}(t)$  совпадает с дифференцированием  $d/d\theta$  по геодезическому параметру  $\theta = \theta^\gamma$  исходной инвариантной связности  ${}^\gamma\nabla$ ,  $\gamma \in \mathbb{R}$ , на  $\text{Car}_2$ . Тогда в этой же связности  ${}^\gamma\nabla$  на  $\text{Car}_4 \supset \text{Car}_2 \otimes \text{Car}_2$  перенос касательных векторов  $(W)_{\mathbf{r}(t(\theta))}$  задается формулой

$$({}^\gamma\nabla_W W)_{\mathbf{r}(t)} = -2\gamma(1-2t)[t(1-t)]^{-2\gamma} \{ (1-t)X_1 + tX_4 \}_{\mathbf{r}(t)}. \quad (13.8)$$

**Доказательство.** По свойству символа  $\nabla$  из (13.7) следует

$$\begin{aligned} \nabla_W W &= \nabla_W [t(1-t)]^{-\gamma} (X_1 - X_4) = \\ &= \{ W [t(1-t)]^{-\gamma} \} (X_1 - X_4) + [t(1-t)]^{-\gamma} \nabla_W (X_1 - X_4), \end{aligned} \quad (13.9)$$

где для упрощения обозначений опущено, что рассмотрение ведется только на кривой  $\mathbf{r}(t)$ . Связь между дифференцированием  $W = d/d\theta$  вдоль кривой и дифференцированием  $d/dt$  навязывается их связью (12.6) в  $\text{Car}_2$ . Поэтому

$$\begin{aligned} \{ W [t(1-t)]^{-\gamma} \} (X_1 - X_4) &= \left\{ [t(1-t)]^{-\gamma+1} \frac{d}{dt} [t(1-t)]^{-\gamma} \right\} (X_1 - X_4) = \\ &= -\gamma [t(1-t)]^{-2\gamma} (1-2t) (X_1 - X_4). \end{aligned} \quad (13.10)$$

Преобразуем теперь второе слагаемое:

$$\begin{aligned} [t(1-t)]^{-\gamma} \nabla_W (X_1 - X_4) &= [t(1-t)]^{-2\gamma} \nabla_{(X_1 - X_4)} (X_1 - X_4) = \\ &= [t(1-t)]^{-2\gamma} \{ \nabla_{X_1} X_1 + \nabla_{X_4} X_4 - \nabla_{X_1} X_4 - \nabla_{X_4} X_1 \} = \\ &= \gamma [t(1-t)]^{-2\gamma} \{ [1 - 2r_1(t)]X_1 + [1 - 2r_4(t)]X_4 - 2[r_4(t)X_1 + r_1(t)X_4] \} = \\ &= -\gamma [t(1-t)]^{-2\gamma} (1-2t)^2 (X_1 + X_4). \end{aligned} \quad (13.11)$$

Подставляя (13.10) и (13.11) в (13.9), получаем

$$\nabla_W W = -\gamma[t(1-t)]^{-\gamma}(1-2t)\{W - [t(1-t)]^{-\gamma}(1-2t)(X_1 + X_4)\}. \quad (13.12)$$

Отсюда, приведя подобные члены, приходим к (13.8).

СЛЕДСТВИЕ. Единственным кандидатом на наследственную инвариантную связность является связность  ${}^0\nabla$ ,  $\gamma = 0$ .

В самом деле, векторные поля  $X_1$  и  $X_4$  линейно независимы, поэтому при  $0 < t < 1$  вектор  $(1-t)X_1 + tX_4 \neq 0$ . Значит, при  $\gamma \neq 0$  правая часть (13.8) обращается в нуль только при  $t = 1/2$ . Это означает, что поле касательных векторов к кривой  $\{\mathbf{r}(t(\theta))\}$  не параллельно вдоль  $\{\mathbf{r}(t(\theta))\}$ . При  $\gamma = 0$  правая часть (13.8) равна нулю тождественно.

Как утверждает теорема 21.1 из [13], если  $\{P_s\}$  есть экспонентное семейство распределений с каноническим параметром  $\mathbf{s}$ , т.е. оно является вполне геодезической поверхностью в инвариантной плоской линейной связности, отвечающей  $\gamma = 0$ , с геодезической параметризацией  $\mathbf{s}$ , то таково же семейство степеней  $\{P_S^N\}$  при любом  $N \in \mathbb{N}$ . Отсюда вытекает (см. [5])

ТЕОРЕМА 13.2. *Логарифмическая плоская связность  ${}^0\nabla$  является единственной инвариантной относительно категории марковских отображений линейной связностью, наследственной при переходе к степеням распределений.*

Напомним, что только эта связность является наследственной также относительно переходов к условным распределениям; см. [13, теорема 12.3].

**14.** Опишем в явном виде геодезическую линию в связности  ${}^\gamma\nabla$  с  $\gamma \neq 0$ , проходящую по  $\text{Int Cap}(\Omega_n, \mathcal{A}_n)$ .

ТЕОРЕМА 14.1. *Семейство векторов  $\mathbf{p}(\theta)$*

$$p_j(\theta) = [\theta q_j^\gamma + (1-\theta)r_j^\gamma]^{1/\gamma} [H(\theta)]^{-1}, \quad j = 1, \dots, n, \quad (14.1)$$

$$H(\theta) = \sum_{k=1}^n [\theta q_k^\gamma + (1-\theta)r_k^\gamma]^{1/\gamma}, \quad (14.2)$$

отвечает геодезической  $\{P_\theta(\cdot)\}$ , проходящей через  $R = P_0$  и  $Q = P_1, Q, R \in \text{Int Cap}(\Omega_n, \mathcal{A}_n)$ . Отрезок  $[\theta^-, \theta^+]$  изменения параметра  $\theta$ ,  $\theta^- < 0 < 1 < \theta^+$ , определяется пересечением условий

$$\theta q_j^\gamma + (1-\theta)r_j^\gamma \geq 0, \quad j = 1, \dots, n. \quad (14.3)$$

Параметр  $\theta$  не является, вообще говоря, каноническим. Геодезический параметр  $z$  определяется квадратурой

$$dz = C[H(\theta)]^{-2\gamma} d\theta \quad \forall C \in \mathbb{R} - \{0\}. \quad (14.4)$$

Доказательство. Продифференцировав (14.1), имеем

$$p_j'(\theta) = [\xi_j(\theta) - \eta(\theta)]p_j(\theta), \quad (14.5)$$

$$\xi_j(\theta) = \frac{1}{\gamma}(q_j^\gamma - r_j^\gamma)\{\theta q_j^\gamma + (1-\theta)r_j^\gamma\}^{-1}, \quad (14.6)$$

$$\eta(\theta) = H'(\theta)[H(\theta)]^{-1} = \frac{d}{d\theta} \ln H(\theta). \quad (14.7)$$

При дифференцировании тождественного равенства  $\sum_{j=1}^n p_j(\theta) = 1$  получаем  $\sum_{j=1}^n p'_j(\theta) = 0$ , откуда, суммируя правую часть (14.5) по  $j$ , находим

$$\sum_{j=1}^n \xi_j(\theta) p_j(\theta) = \eta(\theta). \quad (14.8)$$

Отсюда для семейства векторов  $(V)_{\mathbf{p}} = d/d\theta$  вдоль кривой (14.1) при  $c = -\eta(\theta)$  по (11.19)

$$V = \sum_{j=1}^n \xi_j(\theta) X_j. \quad (14.9)$$

Получим теперь две необходимые формулы. Дифференцируя (14.6) по параметру, устанавливаем

$$\frac{d\xi_i(\theta)}{d\theta} = -\gamma[\xi_j(\theta)]^2. \quad (14.10)$$

Запишем еще (12.3) и (12.4) единой формулой

$${}^{\gamma}\nabla_{X_k} X_j|_{\mathbf{p}} = \gamma\delta_{jk} X_j|_{\mathbf{p}} - \gamma[p_k X_j + p_j X_k]|_{\mathbf{p}}. \quad (14.11)$$

Перейдем к вычислению переноса:

$$\begin{aligned} {}^{\gamma}\nabla_V V &= {}^{\gamma}\nabla_V \left[ \sum_{j=1}^n \xi_j(\theta) X_j \right] = \sum_{j=1}^n [V \xi_j(\theta)] X_j + \sum_{j=1}^n \xi_j(\theta) \cdot {}^{\gamma}\nabla_V X_j = \\ &= \sum_{j=1}^n \left[ \frac{d}{d\theta} \xi_j(\theta) \right] X_j + \sum_{j,k=1}^n \xi_j(\theta) \xi_k(\theta) \cdot {}^{\gamma}\nabla_{X_k} X_j = \\ &= -\gamma \sum_{j=1}^n [\xi_j(\theta)]^2 X_j + \gamma \sum_{j=1}^n [\xi_j(\theta)]^2 X_j - 2\gamma \sum_{j,k=1}^n \xi_k(\theta) p_k(\theta) \xi_j(\theta) X_j = -2\gamma\eta(\theta)V, \end{aligned}$$

где в ходе выкладок мы воспользовались (14.8) и (14.9). Таким образом, при бесконечно малом переносе поля касательных векторов вдоль параметризованной кривой (14.1) оно меняется, оставаясь пропорциональным себе,

$${}^{\gamma}\nabla_V V = -2\gamma\eta(\theta)V. \quad (14.12)$$

Касательный вектор в параметризации (14.4) пропорционален исходному, а именно

$$W = \frac{d}{dz} = C[H(\theta)]^{2\gamma} \frac{d}{d\theta} = C[H(\theta)]^{2\gamma} V. \quad (14.13)$$

Отсюда

$$\begin{aligned} {}^{\gamma}\nabla_W W &= C^2 [H(\theta)]^{4\gamma} \cdot \nabla_V V + 2\gamma C^2 [H(\theta)]^{4\gamma-1} H'(\theta) V = \\ &= C^2 [H(\theta)]^{4\gamma} \{ -2\gamma H'(\theta) [H(\theta)]^{-1} + 2\gamma H'(\theta) [H(\theta)]^{-1} \} V = 0, \end{aligned}$$

что и требовалось показать.

В несколько иной форме сходные рассуждения были проведены Амары [14]. При  $\gamma = 0$  формулы (14.1) и (14.2), а также формула (14.6), содержат неопределенности, которые легко раскрыть (см. [14]) и получить представление классического экспонентного семейства в канонической параметризации  $z$ , совпадающей в этом случае с  $\theta$  (см. (14.4)). Мы отсылаем за точными формулами к [13, § 18], обращая внимание лишь на то, что  $\theta^\pm = \pm\infty$  при  $\gamma = 0$ . Параметр  $\theta$  будет каноническим также и при  $\gamma = 1$ , ибо тогда  $H(\theta) = 1$  тождественно.

Еще один легко интегрируемый случай отвечает римановой геометрии,  $\gamma = 1/2$ ,

$$d\sigma = \frac{d\theta}{1 - 2\theta(1 - \Sigma) + 2\theta^2(1 - \Sigma)}, \quad (14.14)$$

$$\Sigma = \sum_{k=1}^n \sqrt{r_k q_k}; \quad (14.15)$$

ср. (12.11) и (12.12).

**15.** Самостоятельный интерес представляет описание геодезической  $\{P_\theta\}$ , проходящей через  $R \in \text{Int Cap}(\Omega_n, \mathcal{A}_n)$  в данном направлении.

ТЕОРЕМА 15.1. Семейство векторов  $\mathbf{p}(\vartheta)$

$$\mathbf{p}_j(\vartheta) = r_j(1 + \gamma\vartheta u_j)^{1/\gamma} [\mathcal{H}(\vartheta)]^{-1}, \quad j = 1, \dots, n, \quad (15.1)$$

$$\mathcal{H}(\vartheta) = \sum_{k=1}^n r_k(1 + \gamma\vartheta u_k)^{1/\gamma}, \quad (15.2)$$

отвечает геодезической  $\{P_\theta\}$ , проходящей через  $R = P_0 \in \text{Int Cap}(\Omega_n, \mathcal{A}_n)$ , по направлению

$$\mathbf{U} = \frac{d \ln \mathbf{p}(\vartheta)}{d\vartheta}, \quad \sum_{k=1}^n u_k r_k = 0. \quad (15.3)$$

Отрезок  $[\vartheta^-, \vartheta^+]$  изменения параметра  $\vartheta$ ,  $\vartheta^- < 0 < \vartheta^+$ , определяется пересечением условий

$$1 + \gamma\vartheta u_j \geq 0, \quad j = 1, \dots, n. \quad (15.4)$$

Параметр  $\vartheta$  не является, вообще говоря, каноническим (ср. теорему 14.1). Умножение вектора  $\mathbf{u}$  на число приводит к той же геодезической, с пропорциональным изменением геодезического параметра. Последний по-прежнему связан с  $\vartheta$  через нормирующий делитель  $\mathcal{H}(\vartheta)$  аналогом (14.4):

$$dz = C[\mathcal{H}(\vartheta)]^{-2\gamma} d\vartheta. \quad (15.5)$$

ЛЕММА 15.2. Пусть  $\vartheta^- < \vartheta_0 < \vartheta^+$ ,  $\vartheta_0 \neq 0$ ,  $\mathbf{q} = \mathbf{p}(\vartheta_0)$ . Тогда семейство  $\{\mathbf{p}(\vartheta)\}$ , см. (15.1) описывается семейством (14.1), т.е.

$$\mathbf{p}(\vartheta) = \mathbf{p}(\theta), \quad (15.6)$$

с дробно-линейной связью параметров  $\theta$  и  $\vartheta$ ,

$$\theta = \vartheta[\mathcal{H}(\vartheta_0)]^\gamma \{\vartheta_0 - \vartheta + \vartheta[\mathcal{H}(\vartheta_0)]^\gamma\}^{-1}, \quad (15.7)$$

$$H(\theta) = \mathcal{H}(\vartheta)\vartheta_0^{1/\gamma} \{\vartheta_0 - \vartheta + \vartheta[\mathcal{H}(\vartheta_0)]^\gamma\}^{-1/\gamma}. \quad (15.8)$$

Доказательство. При  $\vartheta = \vartheta_0$  из (15.1) имеем

$$q_j = r_j(1 + \gamma\vartheta_0 u_j)^{1/\gamma} \mathcal{H}_0^{-1} \quad \forall j,$$

где  $\mathcal{H}_0 = \mathcal{H}(\vartheta_0)$ . Отсюда можно выразить  $u_j$ :

$$u_j = \frac{1}{\gamma\vartheta_0} \left[ \frac{q_j^\gamma}{r_j^\gamma} \mathcal{H}_0^\gamma - 1 \right] \quad \forall j. \quad (15.9)$$

Подставляя (15.9) в (15.1), после преобразований получаем при любом  $\vartheta \in [\vartheta^-, \vartheta^+]$

$$\begin{aligned} p_j(\vartheta) &= \left\{ r_j^\gamma \left( 1 - \frac{\vartheta}{\vartheta_0} \right) + \frac{\vartheta}{\vartheta_0} \mathcal{H}_0^\gamma q_j^\gamma \right\}^{1/\gamma} [\mathcal{H}(\vartheta)]^{-1} = \\ &= \left\{ r_j^\gamma \frac{\vartheta_0 - \vartheta}{\mathcal{D}(\vartheta, \vartheta_0)} + q_j^\gamma \frac{\vartheta \mathcal{H}_0^\gamma}{\mathcal{D}(\vartheta, \vartheta_0)} \right\}^{1/\gamma} [\mathcal{D}(\vartheta, \vartheta_0)]^{1/\gamma} [\mathcal{H}(\vartheta)]^{-1} \vartheta_0^{-1/\gamma} = \\ &= \{ r_j^\gamma (1 - \theta) + q_j^\gamma \theta \}^{1/\gamma} [H(\theta)]^{-1} = p_j(\theta), \end{aligned}$$

где для краткости обозначено

$$\mathcal{D}(\vartheta, \vartheta_0) = \vartheta_0 - \vartheta - \vartheta \mathcal{H}_0^\gamma \quad (15.10)$$

и коэффициент  $\theta$  при  $q_j^\gamma$  связан с  $\vartheta$  по (15.7), а нормирующий делитель  $H(\theta)$  по единственности выражается через  $\mathcal{H}(\vartheta)$  посредством (15.8). Этим лемма полностью доказана, откуда следует основное утверждение теоремы. Остается проверить формулу (15.5). Для этого продифференцируем (15.7), откуда

$$d\theta = \mathcal{H}_0^\gamma \vartheta_0 [\mathcal{D}(\vartheta, \vartheta_0)]^{-2} d\vartheta. \quad (15.11)$$

Подставим теперь (15.8) и (15.11) в выражение (14.4) для геодезического параметра:

$$\begin{aligned} dz &= C[H(\theta)]^{-2\gamma} d\theta = C\vartheta_0^2 [\mathcal{D}(\vartheta, \vartheta_0)]^2 [\mathcal{H}(\vartheta)]^{-2\gamma} \mathcal{H}_0^\gamma \vartheta_0 [\mathcal{D}(\vartheta, \vartheta_0)]^{-2} d\vartheta = \\ &= \{ C\vartheta_0^{-1} [\mathcal{H}(\vartheta_0)]^\gamma \} [\mathcal{H}(\vartheta)]^{-2\gamma} d\vartheta, \end{aligned}$$

что и требовалось показать.

Интересно было бы выяснить смысл величин  $H(\theta)$  и  $\mathcal{H}(\vartheta)$  как инвариантов; ср. [4], также выше п. 6.

**16.** Так как и кручение, и кривизна являются тензорными полями, то достаточно вычислить их на базисных полях. По определению кручения (см. [23, гл. 1, § 8])

$$T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]. \quad (16.1)$$

Операторы инвариантной связности симметричны (см. (12.3)), а базисные дифференцирования, согласно (11.21), коммутируют. Отсюда при  $j \neq k$

$$T(X_j, X_k) = -\gamma(p_j X_k + p_k X_j) + \gamma(p_k X_j + p_j X_k) = 0.$$

При  $j = k$   $T(X_j, X_j) = 0$  в силу кососимметричности определения (16.1). Отсюда вытекает (см. [5])

**ТЕОРЕМА 16.1.** *Кручение любой инвариантной связности тождественно равно нулю,*

$$\gamma \nabla_X Y - \gamma \nabla_Y X - XY + YX = 0. \quad (16.2)$$

Прежде чем переходить к вычислению кривизны, построим явные выражения действия базисных полей  $X_k$  на базисные функции  $p_j = p_j(\mathbf{p})$ , где  $\mathbf{p} = \langle p_1, \dots, p_k \rangle$ .

**ЛЕММА 16.2.** *Действие базисного поля  $X_k$  на базисную функцию  $p_j$  есть*

$$X_k p_j = -p_k p_j, \quad j \neq k; \quad X_k p_k = p_k - p_k^2,$$

что можно записать единой формулой

$$X_k p_j = \delta_{jk} p_k - p_k p_j. \quad (16.3)$$

**Доказательство.** В силу симметрии достаточно рассмотреть случай  $k = 1$ . Согласно (11.6) уравнение чевианы первого семейства имеет вид

$$p_1(t) = t, \quad p_j(t) = q_j(1 - t), \quad j \geq 2,$$

где  $t$  — натуральный параметр вдоль чевианы, а  $X = t(1 - t) \frac{d}{dt}$  (см. (12.6)). Отсюда

$$X_1 p_1 = t(1 - t) = p_1 - p_1^2, \quad X_1 p_j = q_j(1 - t)t = -p_1 p_j, \quad j \geq 2,$$

что и требовалось доказать.

Рассмотрим теперь оператор кривизны

$$R(X, Y) = \nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X, Y]}, \quad (16.4)$$

и его значения на базисных полях  $X_j$

$$R(X_i, X_j) X_k = R_{kij}^l X_l \quad \forall k, i, j, \quad (16.5)$$

где по индексу  $l$  подразумевается суммирование.

**ТЕОРЕМА 16.2.** *Для инвариантной связности  $\gamma \nabla$  тензор кривизны  $R_{kij}^l$  задается формулой*

$$R(X_i, X_j) X_k = \gamma(\gamma - 1) p_k [(p_j - \delta_{jk}) X_i - (p_i - \delta_{ik}) X_j]. \quad (16.6)$$

**СЛЕДСТВИЕ.** *Тензор  $\gamma R$  кривизны связности  $\gamma \nabla$  пропорционален тензору  ${}^{1/2}R$  кривизны связности  ${}^{1/2}\nabla$ , отвечающей инвариантной римановой метрике,*

$$\gamma R = 4\gamma(1 - \gamma) \cdot {}^{1/2}R. \quad (16.7)$$

**Доказательство.** Так как по (11.21) базисные поля  $X_i$  и  $X_j$  коммутируют, нам достаточно вычислить  $\gamma \nabla_{X_i} \gamma \nabla_{X_j} X_k$ . Для унификации выкладки вместо выражений (12.3) и (12.4) воспользуемся единой формулой (14.11):

$$\gamma \nabla_{X_i} X_j = \gamma \delta_{ij} X_j - \gamma(p_i X_j + p_j X_i). \quad (16.8)$$

Последовательно вычисляем

$$\begin{aligned}\nabla_{X_j} X_k &= \gamma \delta_{jk} X_k - \gamma p_j X_k - \gamma p_k X_j, \\ \nabla_{X_i} \gamma \delta_{jk} X_k &= \underline{\gamma^2 \delta_{ik} \delta_{jk} X_k} - \underline{\gamma^2 \delta_{jk} p_i X_k} - \gamma^2 \delta_{jk} p_k X_i, \\ -\nabla_{X_j} \gamma p_j X_k &= -\underline{\gamma(\delta_{ij} p_i - p_i p_j) X_k} - \underline{\gamma^2 p_j \delta_{ik} X_k} + \underline{\gamma^2 p_j p_i X_k} - \gamma^2 p_j p_k X_i, \\ -\nabla_{X_i} \gamma p_k X_j &= -\gamma(\delta_{ik} p_i - p_i p_k) X_j - \underline{\gamma^2 p_k \delta_{ij} X_j} + \underline{\gamma^2 p_k (p_i X_j + p_j X_i)},\end{aligned}$$

где подчеркнутые одной чертой слагаемые симметричны при перестановке  $i$  и  $j$ . Сложив последние три равенства, получаем

$$\nabla_{X_i} \nabla_{X_j} X_k = -\gamma^2 \delta_{jk} p_k X_i + \gamma^2 p_j p_k X_i - \gamma(\delta_{ik} p_i - p_i p_k) X_j + \dots,$$

где многоточием обозначены пропущенные симметричные слагаемые. При этом учтено, что два слагаемых, подчеркнутые двойной чертой, имеют симметричную сумму. Окончательно,

$$\nabla_{X_i} \nabla_{X_j} X_k = -\gamma p_k [\gamma(\delta_{jk} - p_j) X_i + (\delta_{ik} - p_i) X_j] + \dots, \quad (16.9)$$

где использовано, что  $\delta_{ik} p_i = \delta_{ik} p_k$ . Вычитая из (16.9) аналогичное выражение для  $\nabla_{X_j} \nabla_{X_i} X_k$ , получающееся перестановкой индексов  $i$  и  $j$ , приходим к (16.6). Формула (16.7) является очевидным следствием (16.6). Доказательство завершено.

**17.** Теперь мы можем полностью охарактеризовать глобальную и локальную структуру системы геодезических линий и вполне геодезических подмногообразий в  $\gamma \nabla$ -связности.

**ТЕОРЕМА 17.1.** Система геодезических линий и вполне геодезических подмногообразий открытого симплекса  $\Sigma^{(1)} = \text{Int Cap}_n$

$$\{\mathbf{p} : p_1 + \dots + p_n = 1; p_j > 0 \quad \forall j\} \quad (17.1)$$

в  $\gamma \nabla$ -связности задается пересечением решетки всех линейных подпространств пространства  $\mathbb{R}^n \ni (r_1, \dots, r_n)$  с лежащей в его открытом положительном ортанте гиперповерхностью

$$\Sigma^{(\gamma)} = \{\mathbf{r} : r_1^{1/\gamma} + \dots + r_n^{1/\gamma} = 1; r_j > 0 \quad \forall j\} \quad (17.2)$$

и последующим гомеоморфизмом  $F_\gamma : \Sigma^{(\gamma)} \rightarrow \Sigma^{(1)}$ ,

$$(r_1, \dots, r_n) \rightarrow (p_1, \dots, p_n) = (r_1^{1/\gamma}, \dots, r_n^{1/\gamma}). \quad (17.3)$$

В каждой точке  $\mathbf{p} \in \Sigma^{(1)}$  отображение  $E_{\text{хр}}$  системы всех линейных подпространств касательного пространства  $T_{\mathbf{p}}$  с естественными (см. (15.3)) координатами

$$(u_1, \dots, u_n), \quad u_1 p_1 + \dots + u_n p_n = 0, \quad (17.4)$$

задает систему всех геодезических линий и вполне геодезических многообразий  $\gamma \nabla$ -связности, проходящих через точку  $\mathbf{p}$ .

Доказательство первого утверждения легко вытекает из теоремы 14.1, второго — из формулы (15.1) и первого утверждения.

## СПИСОК ЛИТЕРАТУРЫ

1. Морозова Е.А., Ченцов Н.Н. Структура семейства стационарных состояний квантовой цепи Маркова. — М., 1976. — 55 с. (Препр. / ИПМ АН СССР; № 130). (РЖМат, 1977, 4В294.)
2. Морозова Е.А., Ченцов Н.Н. Алгебра ограниченных гармонических функций счетной цепи Маркова (некоммутативная теория). — М., 1981. — 28 с. (Препр. / ИПМ АН СССР; № 1). (РЖМат, 1981, 8В353.)
3. Морозова Е.А., Ченцов Н.Н. Эквиварианты в категории марковских отображений // Успехи матем. наук. 1985. Т. 40, вып. 2. С. 189–190. (РЖМат, 1986, 2В423)
4. Морозова Е.А., Ченцов Н.Н. Марковские инварианты пар вероятностных законов. — М., 1986. — 27 с. (Препр. / ИПМ АН СССР; № 121). (РЖМат, 1987, 3В12.)
5. Морозова Е.А., Ченцов Н.Н. Инвариантные линейные связности на совокупностях распределений вероятностей. — М., 1988. — 23 с. (Препр. / ИПМ АН СССР; № 167).
6. Наймарк М.А. Спектральные функции симметрического оператора // Изв. АН СССР. Сер. матем. 1940. Т. 4, № 3. С. 277–318.
7. Сарымсаков Т.А. Введение в квантовую теорию вероятностей. — Ташкент: Фан, 1985. — 184 с.
8. Холево А.С. Вероятностные и статистические аспекты квантовой теории. — М.: Наука, 1980. 320 с.
9. Ченцов Н.Н. Категории математической статистики // Докл. АН СССР. 1965. Т. 164, № 3. С. 511–514. (РЖМат, 1966, 2В75.)
10. Ченцов Н.Н. Инвариантные функции потерь в задачах математической статистики // Успехи матем. наук. 1967. Т. 22, вып. 1. С. 178–180. (РЖМат, 1969, 1В122)
11. Ченцов Н.Н. Несимметричное расстояние между распределениями вероятностей, энтропия и теорема Пифагора // Матем. заметки. 1968. Т. 4, вып. 3. С. 323–332. (РЖМат, 1969, 11В7)
12. Ченцов Н.Н. Общая теория статистического вывода: Дис. ... д-ра физ.-мат. наук. — М.: Ин-т прикл. мат. АН СССР, 1968. — 781 с. (РЖМат, 1969, 12В122.)
13. Ченцов Н.Н. Статистические решающие правила и оптимальные выводы. — М.: Наука, 1972. 520 с. (РЖМат, 1973, 4В250К.)
14. Amari Shun-ichi Differential-Geometrical Methods in Statistics // Lect. Notes. Statist. 1985. 28. 290 с.
15. Cencov N.N. Algebraic foundation of mathematical statistics // Math. Operationsforsch. Statist. Ser. Statistics. 1978. V. 9, № 2. P. 267–276.
16. Cencov N.N. Completely positive Markov transition maps, semi-invariant observables, and stationary states. Rep. // 12-th. Europ. Meeting of Statisticians: Varna, 1979.
17. Cencov N.N. Ergodic theorem for Markov maps // Abstracts of Communications of IV USSR–Japan Symposium on Prob. Theory and Math. Statist., Tbilisi. 1982. № 1. P. 162–164.
18. Csiszar I. On topological properties of  $f$ -divergence // Studia Sci. Math. Hungar. 1962. V. 2. P. 329–339.
19. Effros E.G., Stormer E. Positive projections and Jordan structure in operator algebras // Math. scand. 1979. V. 45, № 1. P. 127–138. (РЖМат, 1980, 11В896)
20. Efron B. Defining the curvature of statistical problem (with application to second order efficiency) (with Discussion) // Ann. Statist. 1975. V. 3, № 6. P. 1189–1242. (РЖМат, 1976, 9В127.)
21. Eilenberg S., Mac Lane S. General theory of natural equivalence // Trans. Amer. Math. Soc. 1945. V. 58. P. 231–294.
22. Emch G.G. Mathematical methods in statistical mechanics and quantum field theory. — Wiley, 1972.

23. Helgason S. Differential Geometry and Symmetric Spaces. — Academic Press, 1962. (РЖМат, 1963, 8А379К.)
24. Kraus K. General state changes in quantum theory // Ann. Phys. 1971. V. 64, № 2. P. 331–335.
25. Morozova E.A., Cencov N.N. Stationary matrices of probabilities for stochastic supermatrix // Proceedings of the Third Japan–USSR Symposium on Prob. Theory, Lect. Notes Math. 1986. № 550. P. 379–418.
26. Morozova E.A., Cencov N.N. Markov maps in noncommutative probability theory and mathematical statistics // Abstracts of Fourth Intern. Vilnius Conf. on Prob. Theory and Math. Statist. V. 2, Vilnius, 1985. P. 218–220.
27. Morozova E.A., Cencov N.N. Markov maps in noncommutative probability theory and mathematical statistics // Prob. Theory and Math. Statist., Proceedings of the Fourth Vilnius Conference. Vol. 2. Eds Yu.V. Prohorov et al. — Utrecht: VNU Sci. Press, 1987. — P. 287–310.
28. Morse N., Sacksteder R. Statistical isomorphism // Ann. Math. Statist. 1966. V. 37, № 2. P. 203–214. (РЖМат, 1966, 12В81.)
29. Neumann J. von. Grundlagen der Quantenmechanik. — Berlin, 1932.
30. Stinespring W.F. Positive functions on  $C^*$ -algebras // Proc. Amer. Math. Soc. 1955. V. 6, № 2. P. 211–216.
31. Wald A. Contributions to the theory of statistical estimation and testing hypotheses // Ann. Math. Statist. 1939. V. 10, № 3. P. 299–326.

## ЕСТЕСТВЕННАЯ ГЕОМЕТРИЯ СЕМЕЙСТВ ВЕРОЯТНОСТНЫХ ЗАКОНОВ <sup>\*)</sup>

### СОДЕРЖАНИЕ

- § 0. Введение. Исторические замечания.
  - § 1. Задача статистической точечной оценки как обратная задача теории вероятностей.
  - § 2. Категория статистических решающих правил и эквивалентность статистических экспериментов.
  - § 3. Инварианты пары распределений вероятностей и информационные количества.
  - § 4. Задача различения нескольких простых гипотез.
  - § 5. Аддитивный инвариантный тензор информации Фишера.
  - § 6. Инвариантные линейные связности в многообразиях распределений вероятностей.
  - § 7. Канонические экспоненциальные семейства распределений вероятностей.
  - § 8. Несимметричная пифагорова геометрия информационных количеств.
  - § 9. Параметрическая задача статистического оценивания. Неравенство информации.
  - § 10. Параметрическая задача статистического оценивания. Интегральное неравенство информации.
  - § 11. Параметрическая задача статистического оценивания. Асимптотически оптимальные оценки.
  - § 12. Бесконечномерные квазиоднородные многообразия распределений вероятностей. Информационные поперечники.
  - § 13. Геометризация статистической теории (краткий библиографический обзор).
- Список литературы.  
Именной указатель.  
Предметный указатель.

---

<sup>\*)</sup>Итоги науки и техники. Сер. Современные проблемы математики. Фундаментальные направления. 1991. Т. 83. С. 133–265 (совм. с Е.А. Морозовой). Ввиду большого объема статья приводится в сборнике частично.

## § 0. Введение. Исторические замечания

Последние пятнадцать лет были годами интенсивной геометризации математической статистики и, в первую очередь, теории статистических оценок параметра, где развивался адекватный теории дифференциально-геометрический язык для описания гладких многообразий распределений вероятностей на пространствах выборок (см. [53, 112]). Однако принципиальный вопрос, каким образом возможно задавать естественную дифференцируемую структуру на совокупностях распределений вероятностей, возник много раньше, и не только в связи с потребностями математической статистики (но и теории динамических систем, статистической физики и т.п.). В частности, этот вопрос обсуждался в неопубликованных лекциях А.Н. Колмогорова, прочитанных осенью 1955 г. в институте Анри Пуанкаре. В качестве естественных характеристик различия двух распределений вероятностей на одном и том же измеримом пространстве  $(\Omega, \mathcal{A})$  исходов  $\omega$  А.Н. Колмогоров (см. [101]) рассматривал, в частности, *расстояние по вариации*

$$d_V(P, Q) = \int_{\Omega} |P(d\omega) - Q(d\omega)| \quad (0.1)$$

и “расстояние” Хеллингера

$$d_H(P, Q) = \int_{\Omega} |\sqrt{P(d\omega)} - \sqrt{Q(d\omega)}|^2. \quad (0.2)$$

Однако ни у этих двух характеристик, ни у их очевидного обобщения

$$d_r(P, Q) = \int_{\Omega} |[P(d\omega)]^{1/r} - [Q(d\omega)]^{1/r}|^r, \quad 1 \leq r < \infty, \quad (0.3)$$

непосредственного статистического смысла обнаружено не было.

После того как было доказано неравенство информации и выяснена фундаментальная роль матрицы информации Фишера, возникла идея эксплуатировать на многообразиях распределений вероятностей риманову метрику с фундаментальным тензором, задаваемым матрицей Фишера; см. Рао [179], Джеффрейс [146]. Геометрически эта идея выглядела очень привлекательно. Рассмотрим для наглядности какую-либо *совокупность мультиномиальных распределений*, т.е. совокупность  $\text{Car}_n$  всех распределений вероятностей на конечной алгебре  $\mathcal{A}$  событий, порождаемой атомами  $A_1, \dots, A_n$ . Каждое распределение вероятностей  $P$  на  $\mathcal{A}$  задается строкой  $p_1 = P(A_1), \dots, p_n = P(A_n)$ :

$$\mathbf{p}: \forall i \quad 0 \leq p_i \leq 1; \quad p_1 + \dots + p_n = 1, \quad (0.4)$$

а вся их совокупность отвечает единичному симплексу  $n$ -мерного пространства с вершинами в единичных ортах. Элементу дуги  $ds$  в информационной римановой метрике Фишера отвечает в этих координатах форма

$$ds^2 = \frac{(dp_1)^2}{p_1} + \dots + \frac{(dp_n)^2}{p_n}. \quad (0.5)$$

Другими словами, квадрат дифференциала длины дуги измеряется хи-квадрат величиной.

Если сделать естественную замену координат

$$\forall i \quad z_i = 2\sqrt{p_i}, \quad (0.6)$$

то симплекс превратится в положительный ортант поверхности сферы радиуса 2:

$$z : \forall i \quad 0 \leq z_i \leq 2; \quad z_1^2 + \dots + z_n^2 = 4, \quad (0.7)$$

а фишера форма будет задавать евклидово расстояние на ней:

$$ds^2 = (dz_1)^2 + \dots + (dz_n)^2. \quad (0.8)$$

Однако и здесь попытки найти статистический смысл соответствующему *расстоянию* “в большом”, еще ранее выписанному *Бхаттачария* [80],

$$s_F(P, Q) = 2 \arccos \int_{\Omega} \sqrt{P(d\omega)Q(d\omega)}, \quad (0.9)$$

оказались безуспешными [7, 79, 192]. Оно оказалось интересно только локально (когда оно эквивалентно  $2\sqrt{d_H}$ ). Впрочем, сюда также входит его использование для “бескоординатного” определения дифференцируемости семейства распределений; см. ниже § 9 и приведенную там библиографию. Следует, однако, подчеркнуть, что объем, определяемый на гладком семействе полем фундаментальных тензоров Фишера, как указал Джеффрийс [147, 146], играет важную роль в теории байесовского оценивания; ср. ниже § 10.

Следующая существенная идея состояла в том, чтобы попытаться найти на многообразиях распределений вероятностей имеющую прямой статистический смысл линейную связность. Эта идея была выдвинута первым из авторов настоящей статьи и реализована вторым [26]. Была построена плоская связность, в которой параллельный перенос касательного вектора  $dP$  не зависит от пути переноса. На симплексах (0.4) перенос  $(\mathbf{p}, d\mathbf{p}) \rightarrow (\mathbf{q}, d\mathbf{q})$  описывается правилом

$$\forall i \quad dq_i = q_i \left[ \frac{dp_i}{p_i} - \sum_{k=1}^n \frac{dp_k}{p_k} q_k \right], \quad (0.10)$$

а для бесконечномерных многообразий — его обобщением в терминах производной Радона–Никодима. Геодезическими линиями и вполне геодезическими поверхностями в этой связности (теперь ее принято называть экспоненциальной) являются *канонические экспоненциальные семейства распределений вероятностей с естественной канонической параметризацией* в качестве геодезической, т.е. семейства, задаваемые плотностями по доминирующей мере

$$\frac{p(\omega; s)}{p(\omega; 0)} = \frac{dP_s}{sP_0}(\omega) \exp \left[ \sum_{j=1}^m s^j q_j(\omega) - H(s) \right], \quad (0.11)$$

где  $H(s)$  — *логарифм нормирующего делителя*,

$$H(s) = \ln \int_{\Omega} \exp \left[ \sum_j s^j q_j(\omega) \right] P_0(d\omega) = I(P_s : P_0) - \sum_j s^j \int_{\Omega} q_j(\omega) P_0(d\omega), \quad (0.12)$$

$I(P : P_0)$  — информационное уклонение Кульбака [156, 155],

$$I(P : P_0) = \int_{\Omega} [\ln P_0(d\omega) - \ln P(d\omega)] P_0(d\omega). \quad (0.13)$$

Удобно направляющие статистики  $q_j(\omega)$  выбирать  $P_0$ -несмещенными:

$$\forall j \quad q_j(\omega; 0) = q_j(\omega) - \int_{\Omega} q_j(\omega) P_0(d\omega), \quad (0.14)$$

что равносильно выбору

$$\forall j \quad q_j(\omega, 0) = \left. \frac{\partial}{\partial s^j} \ln p(\omega, \mathbf{s}) \right|_{\mathbf{s}=0}. \quad (0.15)$$

Однако и здесь есть свое “но”, свой разрыв между геометрией и статистикой. Естественная каноническая параметризация “ $s$ ” сама по себе в статистических приложениях не возникает. Зато имеет интерес для статистики двойственная ей натуральная статистическая параметризация, определяемая  $P_s$ -средними от направляющих статистик

$$\mathbf{s} \leftrightarrow \mathbf{t} : \quad \forall j \quad t_j(\mathbf{s}) = \int_{\Omega} q_j(\omega) P_s(d\omega). \quad (0.16)$$

А именно, натуральные статистические параметризации канонических экспоненциальных семейств и только они допускают эффективное оценивание, обрабатывающее неравенство информации в равенство. Для одномерных семейств это обстоятельство было замечено еще Бхаттачария [81].

Информационное количество Кульбака  $I(P : Q)$  определяет асимптотику ошибки при проверке простой гипотезы  $P$  против простой альтернативы  $Q$ . Вообще говоря,  $I(P : Q) \neq I(Q : P)$ , что отражает несимметрию отношения похожести случайных явлений. То, что это же количество Кульбака возникает и в теории оценивания параметров, наводит на мысль, что информационное уклонение  $I(P : Q)$  есть универсальная характеристика “непохожести” распределения вероятностей  $Q$  на  $P$ . При близких альтернативах

$$I(P : Q) \sim [s_F(P, Q)]^2 / 2.$$

Таким образом, количество Кульбака является своего рода несимметричным аналогом половины квадрата расстояния. Как было впервые показано нами в [32], эта аналогия заходит так далеко, что для проектирования на каноническое экспоненциальное семейство (и для двойственной задачи) выполняется несимметричный вариант теоремы Пифагора.

Пусть  $\mathcal{P} = \{P_s\}$  — каноническое экспоненциальное семейство распределений вероятностей и притом регулярное (см. ниже определение 7.2);  $R(\cdot)$  — распределение вероятностей, взаимно абсолютно непрерывное с распределениями  $P_s \in \mathcal{P}$ . Тогда существует и единственно

$$P_{\sigma} = \operatorname{argmin} \mathcal{P} I(P_s : R), \quad (0.17)$$

где значение  $\sigma$  определяется через значение  $\tau = t(\sigma)$  двойственного натурального статистического параметра (0.16) равенством

$$\forall j \quad \tau_j = \int_{\Omega} q_j(\omega) R(d\omega). \quad (0.18)$$

Для любого распределения  $P_s \in \mathcal{P}$

$$I(P_s : R) = I(P_s : P_\sigma) + I(P_\sigma : R). \quad (0.19)$$

Заметим, что если в каноническом экспоненциальном семействе (0.11), хотя бы и нерегулярном, существует распределение  $P_\sigma$ , для которого значение  $\tau = t(\sigma)$  двойственного натурального статистического параметра задается (0.18), то несимметричное равенство Пифагора (0.19) выполнено, и справедливо (0.17).

Класс регулярных экспоненциальных семейств достаточно широк: к нему принадлежат, например, все семейства (0.11) с ограниченными направляющими статистиками  $q_j(\omega)$ . В общем случае можно утверждать лишь, что

$$I(P_s : R) \geq I(P_s : P_\sigma) + I(P_\sigma : R),$$

если (0.17) действительно определяет  $P_\sigma \in \mathcal{P}$ , и т.п.

Равенство (0.19), неоднократно переоткрывавшееся и обобщавшееся позднее (см. [49]) другими авторами, было положено нами в [34] в основу информационной теории счетно-параметрического оценивания (т.е. плотности). Двойственная задача состоит в двойственном проектировании распределения  $P_0$  на “ $m$ -коплоскость”  $\mathcal{R}$  распределений  $R(\cdot)$  с фиксированным набором  $\tau_1, \dots, \tau_m$  значений интегралов (0.18):

$$P = \arg \min \mathcal{P} I(P_0 : R). \quad (0.20)$$

Эта задача впервые изучалась (при  $m = 1$ ) в связи с приложениями к статистической физике [23]. Она играет важную роль в теории больших уклонений, что впервые было отмечено для дискретных схем И.Н. Сановым [20] и затем Кульбаком [155]. Ее решение в регулярном случае дает  $P = P_\sigma$ , и для троек  $(P_0, P_\sigma, R)$  выполнено равенство (0.19) с  $P_s = P_0$ ,  $P_\sigma = P$  (см. [34, 110]). Позднее теория двойственности была обобщена Амари [49] и разработана еще далее Лауритценом (см. [160]).

Если принять, что целью статистика в задачах параметрического оценивания является не значение  $\theta$  параметра наблюдаемого распределения вероятностей  $P_\theta$ , а само распределение  $P_\theta(\cdot)$ , и избрать величину  $2\mathbf{I}(P_\theta : P^*)$  в качестве функции потерь при оценке  $P^*(\cdot)$ , то (см. [34, 35, 99]) для любого гладкого семейства  $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$  средний по семейству  $\mathcal{P}$  (в смысле упоминавшегося выше информационного объема Фишера–Джеффриса) риск  $R_{III(N)}$  для любого решающего правила  $III(N)$  оценивания по  $N$  независимым наблюдениям не может быть меньше

$$\frac{1}{N} \dim \Theta + K(N),$$

где величина поправки  $K(N) = o(N^{-1})$  может быть грубо указана. В то же время риск  $R_{II(N)}(P_\theta)$  оценки наибольшего правдоподобия  $II(N)$  для семейств

ва  $\mathcal{P}$  при любом  $\theta \in \Theta$  равен

$$\frac{1}{N} \dim \Theta + o\left(\frac{1}{N}\right),$$

т.е. оценка  $\Pi(N)$  является асимптотически оптимальной. Эти два очень просто формулируемых результата (см. (99) и ниже § 11) являются, в сущности, квинтэссенцией стандартной теории оценок параметра.

В настоящее время наибольшее внимание привлекают приложения дифференциально-геометрических методов именно в теории оценивания параметров, в первую очередь получение достижимых границ точности оценивания, в частности, асимптотическое уточнение поправки  $K(N)$ , и т.п. В этой статье мы демонстрируем, что геометрические методы могут с успехом использоваться и в других разделах статистической теории, таких, как проверка гипотез и оценивание плотности, когда приходится решать задачу статистического точечного оценивания распределения вероятностей случайной величины по ее независимым наблюдениям; ср. [165].

Рассмотрение с единых позиций широкого класса типов задач позволяет отчетливее уяснить особенности естественного геометрического языка и возможности его адекватного использования. При этом приходится иметь дело с несимметричным обобщением элементарной геометрии Пифагора–Евклида в теории проверки простых гипотез, дифференциальной геометрией многообразий в двух сопряженных линейных связностях — в теории параметрического оценивания, несимметричным обобщением теории емкости и поперечников Колмогорова в задачах оценивания плотностей. Правда, в единообразной постановке мы получаем только главные члены асимптотики убывания минимального риска при естественной геометрической функции потерь.

Задача статистического точечного оценивания является, грубо говоря, обратной задачей теории вероятностей и, как обычно для нетривиальных задач математической физики, она некорректно поставлена (см. [37–39]). Однако она допускает состоятельное решение, если априори задано достаточно узкое семейство  $\mathcal{P}$  распределений вероятностей, к которому “наблюдаемое” распределение  $P$  обязательно принадлежит. При этом максимальная достижимая точность решения зависит от геометрии семейства  $\mathcal{P}$ . Именно ввиду последнего обстоятельства геометрический язык естественно используется для формулировки статистических закономерностей.

Из сказанного вытекает и второе соображение — естественный геометрический язык должен быть согласован со статистическими правилами обработки данных. Однако надо заметить, вслед за [153], что это обстоятельство не всегда учитывается. В настоящей статье основное внимание уделяется тем понятиям геометрии многообразий распределений вероятности, которые инвариантны (эквивариантны) относительно категории статистических решающих правил (иначе называемых марковскими отображениями или марковскими морфизмами) и наследственны относительно ее функтора умножения. Эта категория была введена нами [27] и независимо Морсом и Сакстедером [169] для алгебраической формализации основополагающих идей А. Вальда [195] и Д. Блекуэлла [84, 85] в теории статистических решений и статистических экспериментов; см. также

подходы Ле Кама [162]. С позиций категорного подхода удается понять, какие геометрические понятия “работают” в статистике (как, например, тензор информации Фишера), а какие играют только вспомогательную роль (как риманова метрика, определяемая этим тензором).

Если быть строгим, то в математической статистике приходится иметь дело с дифференциальной геометрией поверхностей в бесконечномерных многообразиях распределений вероятностей. Единственное, по существу, исключение — искривленные экспоненциальные семейства — подсемейства канонических экспоненциальных семейств (0.11), — благодарный объект исследований с применением дифференциально-геометрической техники; ср. [115, 54]. Но общепринятой бесконечномерной дифференциальной геометрии пока не существует. И далеко не все предложения конечномерной теории безоговорочно сохраняются в бесконечномерной. Даже теория геодезических таит здесь сюрпризы; см. ниже § 6 и § 7. Поэтому особое внимание мы уделяем достаточным условиям, обеспечивающим регулярность тех или иных конструкций, особенно в задаче оценивания плотности.

Объем статьи заставляет предполагать у читателя знакомство с основными понятиями дифференциальной геометрии, скажем, в объеме вводных параграфов монографий [34] или [49] или в объеме вводной статьи [113]. В обзоре дополнительно кратко прореферированы основные новые дифференциально-геометрические понятия, возникшие в связи с приложениями к задачам статистики. По-видимому, они представляют интерес и для теоретической физики.

Впрочем, определенный интерес для теоретической физики составляет и основное содержание статьи, особенно в связи с задачами квантовой теории измерения; см. монографию [24] и статью А.С. Холево в настоящем томе, а также [144, 143]. Последняя занимается решениями обратных задач некоммутативной теории вероятностей, описывающей случайные явления микромира, подобно тому как в математической статистике решаются обратные задачи классической теории вероятностей Лапласа–Колмогорова. Оказывается, что некоторые результаты настоящей статьи без особых затруднений переносятся на некоммутативный случай. За более детальным изложением аспектов возникающей теории мы отсылаем к нашим работам [11] и [168].

Рамки настоящей статьи не позволяют дать систематический обзор современного этапа геометризации математической статистики. За деталями мы отсылаем к недавней монографии Амари [49], сборникам [53] и [112]. Специалистам по теории параметрического статистического оценивания можно рекомендовать вводящую в геометрию асимптотических выводов статью Касса [153] (с комментариями), не использующую дифференциально-геометрической техники. Отметим, что наша статья во многих отношениях “дополнительна” к [153], поскольку основное внимание уделяется статистическому смыслу развиваемых концепций, а не геометрической стороне вопроса. Авторы пользуются случаем выразить Ш.-И. Амари, О.Е. Барндорф-Нильсену, П. Блэсиду, С.Т. Дж. Додсону и Р.З. Хасьминскому благодарность за внимание. Основные концепции настоящего обзора излагались нами на недавних семинарах по геометрическим методам математической статистики в Математическом институте им. В.А. Стек-

лова АН СССР. Авторы признательны его руководителям и участникам, в первую очередь А.Т. Фоменко, А.С. Холево, Д.М. Чибисову и А.Н. Ширяеву за их вопросы и замечания, учтенные при подготовке статьи.

\* \* \*

### **§ 13. Геометризация статистической теории (краткий библиографический обзор)**

В настоящее время разработка геометрических подходов в математической статистике и исследование геометрических аспектов статистической теории интенсивно развивается по целому ряду направлений, в первую очередь связанных с задачами параметрического оценивания. Эти исследования не только обогатили арсенал статистических методов, но и, в свою очередь, индуцировали дальнейшее развитие аппарата самой дифференциальной геометрии, которое может оказаться интересным в связи с приложениями не только к статистике, но и к физике. Работы в этом направлении по теории стрингов (струн, но не в смысле хромодинамики) были начаты О.Е. Барндорф-Нильсеном и продолжаются им совместно с коллективом сотрудников [58, 61, 64, 68–71, 88–91]. Сюда же примыкают их работы по (средне)ожидаемой и наблюдаемой геометрии, диктуемым (см. § 11) функцией правдоподобия и ее обобщениями [57, 74, 86–88, 90, 167]; ср. также [154, 160, 166]. Более подробно с ними можно познакомиться по программным статьям [59–69, 64, 73] и монографиям [63, 72]; ср. [125, 126].

За последние двадцать лет получили большое развитие исследования по определению естественных пределов применимости неравенства информации. Кроме основополагающих работ [106, 111, 129, 179], мы упомянули здесь лишь некоторые [41, 119, 142, 180, 197, 198]; см. более подробную библиографию этих работ в [198]. Ссылки на более ранние исследования, включая доказательства Рао и Ченцова неравенства информации для оценок, являющихся дифференцируемыми функционалами от эмпирической функции распределения, даны в [34].

Важный этап в развитии геометростатистики (впервые этот термин был употреблен А.Н. Колмогоровым по отношению к предмету монографии [34]) начался с работы Эфрона [115], в которой в теорию асимптотически оптимального статистического оценивания было введено понятие кривизны статистической модели. Дальнейшему развитию этого направления, которое не отражено в нашем обзоре, посвящены [45, 48, 54, 175, 116, 158, 174, 177]; см. также [82] и достаточно подробный обзор этого направления в [153]. Кроме того, здесь необходимо упомянуть более ранние исследования, начатые С.Р. Рао [180]; см. также [183, 1]. Однако, как нам представляется в контексте § 10 (см. выше пример 10), точка в теории эффективных по второму порядку оценок еще не поставлена.

Развитие дифференциально-геометрических методов математической статистики во многом связано с исследованиями Ш.-И. Амари, который объединил в своем лице геометра и специалиста по теории информации. Среди его работ, кроме упоминавшихся в разных местах настоящего обзора, следует отметить [46, 55, 157], связанные с теорией оценивания при мешающих параметрах. Для более подробного ознакомления с его идеями и результатами мы отсылаем к на-

писанным им очень четко и ясно монографии и обзорам [49–52]; см. также [105, 128, 136].

В настоящем обзоре затронуты лишь два аспекта использования оценок максимального правдоподобия в задачах конечно- и счетно-параметрической статистики. Этот вопрос имеет богатую литературу, включая классические работы Р.А. Фишера [120–124]; см. [10, 107, 117, 137, 163]. Геометрические аспекты этого метода, как уже отмечалось, проанализированы в [63]. В изложении теории канонических экспоненциальных семейств вероятностных мер мы следовали систематическому подходу [28, 34]; укажем для сравнения также [56, 92]. Приложение этой теории к задаче оценивания плотностей, описанное выше в § 12, следует идеям [34].

При изложении теории информационного тензора Фишера и индуцируемой им марковски инвариантной римановой геометрии Фишера–Бхаттачария–Рао были упомянуты исходные работы указанных авторов [120–124, 80, 81, 179–182], а также Джеффрейса [146, 147], и лишь несколько последующих работ, в которых впервые обсуждался тот или иной аспект этой геометрии [7, 33, 49, 52, 95, 159, 160, 176]. По аналогичному принципу сделаны ссылки на теорию информационного количества Кульбака [155, 156, 145], см. [19, 20, 22, 29, 32], а также его обобщений [5, 49, 93–95, 102, 108–110, 172, 185].

Исследования по марковской геометрии и теории инвариантов, начатые Н.Н. Ченцовым [27, 29, 30] и Р. Сакстедером [169, 188, 189], были продолжены в [12–15, 32–34, 168, 176, 177, 187]. Более подробно история возникновения понятия эквивалентных статистических экспериментов, начиная с идей А. Вальда, изложена в [34]; см. также [84, 85, 193, 195].

## СПИСОК ЛИТЕРАТУРЫ

1. *Большев Л.Н.* Уточнение неравенства Крамера–Рао // Теория вероятн. и ее примен. 1961. Т. 6, вып. 3. С. 319–326.
2. *Воробьев Н.Н., Фаддеев Д.К.* Континуализация условных вероятностей // Теория вероятн. и ее примен. 1961. Т. 6, вып. 1. С. 116–118.
3. *Гельфанд И.М., Колмогоров А.Н., Яглом А.М.* К общему определению количества информации // Докл. АН СССР. 1956. Т. 111, № 4. С. 745–748.
4. *Дынкин Е.Б.* Необходимые и достаточные статистики // Успехи матем. наук. 1951. Т. 6, вып. 1. С. 66–90.
5. *Каган А.М.* К теории информационного количества Фишера // Докл. АН СССР. 1963. Т. 153, № 3. С. 277–278.
6. *Каган А.М.* Семейства распределений и разделяющие разбиения // Докл. АН СССР. 1963. Т. 153, № 3. С. 522–525.
7. *Козлов В.П.* Емкость множества в пространстве сигналов и риманова метрика // Докл. АН СССР. 1966. Т. 166, № 4. С. 779–782.
8. *Колмогоров А.Н.* Основные понятия теории вероятностей. — М.–Л.: ОНТИ, 1936.
9. *Колмогоров А.Н.* Три подхода к определению понятия “количества информации” // Пробл. передачи информ. 1965. Т. 1, № 1. С. 3–11.
10. *Линник Ю.В., Митрофанова Н.М.* К асимптотике распределения максимального правдоподобия // Докл. АН СССР. 1963. Т. 149, № 3. С. 518–520.
11. *Морозова Е.А., Ченцов Н.Н.* Структура семейства стационарных распределений для некоммукативной цепи Маркова // Мат. модели статист. физ.. — Тюмень: Изд. ТГУ, 1982.

12. Морозова Е.А., Ченцов Н.Н. Эквиварианты в категории марковских отображений // Успехи матем. наук. 1985. Т. 40, вып. 2. С. 189–190.
13. Морозова Е.А., Ченцов Н.Н. Марковские инварианты пар вероятностных законов — М., 1986. — 27 с. — (Препр. / Ин-т прикл. мат. АН СССР; № 121).
14. Морозова Е.А., Ченцов Н.Н. Инвариантные линейные связности на совокупностях распределений вероятностей. — М., 1988. — 23 с. — (Препр. / Ин-т прикл. мат. АН СССР; № 167).
15. Морозова Е.А., Ченцов Н.Н. Марковская инвариантная геометрия на многообразиях состояний // Современ. пробл. матем. Новейшие достижения. ВИНТИ, 1990. Т. 36. С. 69–102.
16. Норден А.П. О парах сопряженных параллельных перенесений в  $n$ -мерных пространствах // Докл. АН СССР. 1945. Т. 49, № 9.
17. Норден А.П. О внутренней геометрии 2-го рода на гиперплоскости аффинного пространства / Приложение к книге: П.А. Широков, А.П. Широков “Аффинная дифференциальная геометрия”. — М.: Физматгиз, 1969.
18. Розенблат-Рот М. Понятие энтропии в теории вероятностей и его применение к теории передачи информации по каналам связи // Теория вероятн. и ее примен. 1964. Т. 9, вып. 2. С. 238–260.
19. Салихов Н.П. Асимптотические свойства вероятностей ошибок критериев для различения нескольких мультиномиальных схем испытаний // Докл. АН СССР. 1973. Т. 209, № 1. С. 54–57.
20. Санов И.Н. О вероятностях больших отклонений случайных величин // Мат. сб. 1957. Т. 42, № 1. С. 11–44
21. Тихомиров В.М. Поперечники множеств в функциональных пространствах и теория наилучших приближений // Успехи матем. наук. 1960. Т. 15, вып. 3. С. 81–120.
22. Фаддеев Д.К. К понятию энтропии конечной вероятностной схемы // Успехи матем. наук. 1956. Т. 11, вып. 1. С. 227–231.
23. Хинчин А.Я. Математические основания статистической механики. — М.–Л.: Гостехиздат, 1943. — 126 с.
24. Холево А.С. Вероятностные и статистические аспекты квантовой теории. — М.: Наука, 1980. — 320 с.
25. Ченцов Н.Н. Оценка неизвестной плотности распределения по наблюдениям // Докл. АН СССР. 1962. Т. 147, № 1. С. 45–48.
26. Ченцов Н.Н. Геометрия “многообразия” распределений вероятностей // Докл. АН СССР. 1964. Т. 158, № 3. С. 543–546.
27. Ченцов Н.Н. Категории математической статистики // Докл. АН СССР. 1965. Т. 164, № 3. С. 511–514.
28. Ченцов Н.Н. К систематической теории экспоненциальных семейств распределений вероятностей // Теория вероятн. и ее примен. 1966. Т. 11, вып. 3. С. 483–494.
29. Ченцов Н.Н. Инфинитезимальные методы математической статистики / Деп. рукоп. — ВИНТИ, 1966.
30. Ченцов Н.Н. Инвариантные функции потерь в задачах математической статистики // Успехи матем. наук. 1967. Т. 22, вып. 1. С. 178–180.
31. Ченцов Н.Н. Об оценке неизвестного среднего многомерного нормального распределения // Теория вероятн. и ее примен. 1967. Т. 12, вып. 4. С. 619–634.
32. Ченцов Н.Н. Несимметричное расстояние между распределениями вероятностей, энтропия и теорема Пифагора // Матем. заметки. 1968. Т. 4, вып. 3. С. 323–332.
33. Ченцов Н.Н. Общая теория статистического вывода // Матем. заметки. 1969. Т. 5, вып. 5. С. 635–648.
34. Ченцов Н.Н. Статистические решающие правила и оптимальные выводы. — М.: Наука, 1972.
35. Ченцов Н.Н. Оптимальные решающие правила в классических задачах математической статистики // Тез. Междун. конф. по теор. вероятн. и матем. статистике. — Вильнюс, 1973. — С. 325–328.

36. Ченцов Н.Н. Неизмеримые подмножества тихоновского куба // Пробл. матем. физ. и вычислительной матем. — М.: Наука, 1977. — С. 316–320.
37. Ченцов Н.Н. О корректности задачи статистического точечного оценивания. — М., 1979. — 26 с. — (Препр / Ин-т прикл. мат. АН СССР; № 193).
38. Ченцов Н.Н. О статистических оценках неизвестного закона распределения вероятностей // Теория вероятн. и ее примен. 1980. Т. 25, вып. 3. С. 647–648.
39. Ченцов Н.Н. О корректности задачи статистического точечного оценивания // Теория вероятн. и ее примен. 1981. Т. 26, вып. 1. С. 15–31.
40. Ченцов Н.Н. Почему  $L_1$ -подход и что за горизонтом / Приложение к книге: Л. Деврой, Л. Дёрфи “Оценка плотности.  $L_1$ -подход”. — М.: Мир, 1988. — С. 348–361.
41. Шалаевский О.В. Краткое доказательство неравенства Крамера–Рао // Теория вероятн. и ее примен. 1961. Т. 6, вып. 3. С. 352–353.
42. Abou-Jaoude S. Conditions nécessaires et suffisantes de convergence  $L_1$  en probabilité de l’histogramme pour une densité // Ann. l’Inst. H. Poincaré. 1976. Ser. B. V. 12. P. 213–231.
43. Amari S.-I. Theory of information space — a geometrical foundation of the analysis of communication system // RAAG Memoirs // 1968. 4. С. 373–418
44. Amari S.-I. Theory of information spaces — a differential geometrical foundation of statistics // POST RAAG Report. 1980. № 106.
45. Amari S.-I. Differential geometry of curved exponential families — curvatures and information loss // Ann. Statist. 1982. V. 10, № 2. P. 357–387.
46. Amari S.-I. Geometrical theory of asymptotic ancillarity and conditional inference // Biometrika. 1982. V. 69, № 1. P. 1–17.
47. Amari S.-I. Differential geometry of statistical inference // K. Ito and Yu.V. Prokhorov (eds.) “Probability Theory and Mathematical Statistics”. — Lect. Notes Math. 1983. V. 1021. P. 26–40.
48. Amari S.-I. Comparisons of asymptotically efficient tests in terms of geometry of statistical structures // Bull. Statist. Inst. 1983. Proc. 44 Session. V. 2. P. 1190–1206.
49. Amari S.-I. Differential-Geometrical Methods in Statistics. / Lect. Notes Statist. 28. — N.Y.: Springer, 1985.
50. Amari S.-I. Differential geometrical theory of statistics // Different. Geometry in Statist. Inference. Hayward. — Calif.: Inst. Math. Statist. 1987. P. 19–94.
51. Amari S.-I. Differential geometry of a parametric family of invertible linear systems — Riemannian metric, dual affine connections and divergence // Math. Syst. Theory. 1987. V. 20, № 1. P. 53–82.
52. Amari S.-I. Dual connection on the Hilbert bundles of statistical models // C.T.J. Dodson (ed.) “Geometrization of Statistical Theory”. — Lancaster: Univ. Lanc. Dept. Math. Publ., 1987. — P. 123–152.
53. Amari S.-I., Barndorff-Nielsen O.E., Kass R.E., Lauritzen S.I., Rao C.R. Differential Geometry in Statistical Inference / IMS Lect. Notes, X. — Hayward, Calif: Inst. Math. Statist., 1987
54. Amari S.-I., Kumon. M. Differential geometry of Edgeworth expansions in curved exponential family // Ann. Inst. Statist. Math. 1983. 35A. P. 1–24.
55. Amari S.-I., Kumon. M. Estimation in the presence of infinitely many nuisance parameters-geometry of estimating functions // Ann. Statist. 1988. V. 16, № 4. P. 1044–1068.
56. Barndorff-Nielsen O.E. Information and exponential families in statist. theory. — N.Y.: Wiley, 1978.
57. Barndorff-Nielsen O.E. Likelihood and observed geometries // Ann. Statist. 1986. V. 14. P. 856–873
58. Barndorff-Nielsen O.E. Strings tensorial combinants and Bartlett adjustments // Proc. Roy. Soc. London. Ser. A. 1986. V. 406. P. 127–137.
59. Barndorff-Nielsen O.E. Differential and integral geometry in statistical inference // Different. geometry in statist. inference Hayward, Calif. // Inst. Math. Statist. 1987. P. 95–161.

60. *Barndorff-Nielsen O.E.* Differential geometry and statistics: Some mathematical aspects // *Indian J. Math.* 1987. V. 29. P. 335–350.
61. *Barndorff-Nielsen O.E.* On some differential geometric concepts and their relations to statistics // C. T. 1. Dodson (ed.) *Geometrization of Statistical Theory* // Lancaster: Univ. Lane. Dept. Math. Publ. 1987. P. 58–90.
62. *Barndorff-Nielsen O.E.* Likelihood, ancillarity and strings // Proc. 1-st World Congress Bernoulli Soc. // *Math. Statist. Probab.* 1987. V. 2. P. 205–213.
63. *Barndorff-Nielsen O.E.* Parametric Statistical Models and Likelihood / *Lect. Notes Statist.* 50. — N.Y.: Springer, 1988.
64. *Barndorff-Nielsen O.E.* Comment on R.E. Kass. The geometry of asymptotic inference // *Statist. Sci.* 1989. V. 4, № 3. P. 222–227.
65. *Barndorff-Nielsen O.E., Blaesild P.* Exponential models with affine dual foliations // *Ann. Statist.* 1983. V. 11 P. 753–769.
66. *Barndorff-Nielsen O.E.* Reproductive exponential families // *Ann. Statist.* 1983. V. 11. P. 770–782.
67. *Barndorff-Nielsen O.E., Blaesild P.* A note the calculation of Bartlett adjustments // *J. Roy. Statist. Soc., ser. B.* 1986. V. 48. P. 353–358.
68. *Barndorff-Nielsen O.E., Blaesild P.* Strings: Mathematical theory and statistical examples // *Proc. Roy. Soc. London. Ser. A.* 1987. V. 411, № 3. P. 155–176.
69. *Barndorff-Nielsen O.E., Blaesild P.* Strings: Contravariant aspect // *Proc. Roy. Soc. London, ser. A.* 1987. V. 411, № 3. P. 421–444.
70. *Barndorff-Nielsen O.E., Blaesild P.* Coordinate-free definition of structurally symmetric derivative strings // *Adv. Appl. Math.* 1988. V. 9, № 1. P. 1–6.
71. *Barndorff-Nielsen O.E., Blaesild P., Mora M.* Generalized higher order differentiation // *Acta Appl. Math.* 1989. V. 16. P. 243–259.
72. *Barndorff-Nielsen O.E., Cox D.R.* *Asymptotic Techniques for Use in Statistics.* — London: Chapman and Hall, 1989.
73. *Barndorff-Nielsen O.E., Blaesild P., Reid N.* The role of differential geometry in statistical theory // *Internat. Statist. Rev.* 1986. V. 54, № 1. P. 83–96.
74. *Barndorff-Nielsen O.E., Jupp P.E.* Differential geometry, profile likelihood,  $L$ -sufficiency and composite transformation models // *Ann. Statist.* 1988. V. 16. P. 1009–1043.
75. *Barndorff-Nielsen O.E., Jupp P.E.* Approximating exponential models // *Ann. Inst. Statist. Math.* 1989. V. 41. P. 247–267.
76. *Basu D.* Statistical information and likelihood // *Sankya.* 1975. V. 37A, № 1. P. 1–71.
77. *Bates D.M., Watts D.G.* Relative curvature measures of non-linearity (with discussion) // *J. Roy. Statist. Soc., ser. B.* 1980. V. 42, № 1. P. 1–25.
78. *Begun J.M., Hall W.J., Huang W.M., Wellner I.A.* Information and asymptotic efficiency in parametric-nonparametric models // *Ann. Statist.* 1983. V. 11. P. 432–452.
79. *Beran R.* Minimum Hellinger distance estimates for parametric models // *Ann. Statist.* 1977. V. 5. P. 445–463.
80. *Bhattacharyya A.* On a measure of divergence between two statistical populations defined by their probability distributions // *Bull. Calcutta Math. Soc.* 1943. V. 35. P. 99–109.
81. *Bhattacharyya A.* On some analogues of the amount of information and their use in statistical estimation // *Sankhya.* 1946. V. 8, № 1. P. 1–14; 1947. V. 8, № 3. P. 201–218.
82. *Bickel P.J., Chibisov D.M., Van Zwet W.R.* On efficiency of first and second order // *Inst. Statist. Rev.* 1981. V. 49. P. 169–175.
83. *Blackwell D.* Conditional expectation and unbiased sequential estimation // *Ann. Math. Statist.* 1947. V. 18, № 2. P. 105–110.
84. *Blackwell D.* Comparison of experiments // Proc. II Berkeley Symp. *Math. Statist. Probab.* 1951. P. 93–102.
85. *Blackwell D.* Equivalent comparisons of experiments // *Ann. Math. Statist.* 1953. V. 24. P. 265–272.
86. *Blaesild P.* *Yokes: Elemental properties with statistical applications* // C.T.J. Dodson, ed. *Geometrization of Statistical Theory.* — Lancaster: Univ. Lane. Dept. Math. Publications, 1987. — P. 193–198.

87. *Blaesild P.* Further analogies between expected and observed geometries of statistical models // Research Report № 160, Dept. Theor. statist., Aarhus Univ., 1987.
88. *Blaesild P.* Yokes and tensor derived from yokes / ibidem. — 1988, № 173.
89. *Blaesild P.* Alternating sums of derivative costrings / ibidem. — 1989, № 187.
90. *Blaesild P.* Yokes: Orthogonal and extended normal coordinates / ibidem. — 1990, № 206.
91. *Blaesild P., Mora M.* Differential strings: basis theory / ibidem. — 1988, № 183.
92. *Brown L.D.* Fundamentals of statistical exponential families with applications in statistical decision theory. — Hayward, Calif.: Inst. Math. Statist. 1986.
93. *Burdea J.* Informative geometry of probability spaces // Exposition Math. 1986. V. 4. P. 347–378.
94. *Burdea J., Rao C.R.* Entropy differential metric, distance and divergence measures in probability spaces: a unified approach // J. Multivar. Anal. 1982. V. 12. P. 575–596.
95. *Burdea J., Rao C.R.* Differential metric in probability spaces // Probab. Math. Statist. 1982. V. 3. P. 115–132.
96. *Burkholder D.* On the order structure of the set of sufficient subfields // Ann. Math. Statist. 1962. V. 33, № 2. P. 569–599.
97. *Campbell L.L.* The relation between information theory and the differential geometry approach to statistics // Inf. Sci. 1985. V. 35. P. 199–210.
98. *Čencov N.N.* Algebraic foundation of mathematical statistics // Math. Operationsforsch. und Statistic. Ser. Statist. 1978. V. 9, № 2. P. 267–276.
99. *Čencov N.N.* On basic concepts of mathematical statistics // Banach Center Publ. 1980. 6. P. 85–94.
100. *Čencov N.N. (Chentsov N.N.)* Statistical Decision Rules and Optimal Inference / Trans. Math. Monographs 53. — Providence, R. I.: Amer. Math. Soc., 1982.
101. *Chentsov N.N.* The unfathomable influence of Kolmogorov // Ann. Statist. 1990. V. 18, № 3. P. 987–998.
102. *Chernoff H.* A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations // Ann. Math. Statist. 1952. V. 23. P. 493–507.
103. *Chernoff H.* Large sample theory: parametric case // Ann. Math. Statist. 1956. V. 27, № 1. P. 1–22.
104. *Chibisov D.M.* On the normal approximation for a certain class of statistics // Proc. VI Berkeley Symp. Math. Statist. Probab. — 1972. 1. P. 153–174.
105. *Cox D.R., Reid N.* Parameter orthogonality and approximate conditional inference (with discussion) // J. Roy. Statist. Soc. Ser. B. 1987. V. 49, № 1. P. 1–39.
106. *Cramér H.* Contribution to the theory of statistical estimation // Skand. Akt. Tidskr. 1946. V. 29. P. 85–94.
107. *Cramér H.* Mathematical Methods of Statistics. — Princeton Univ. Press, 1946.
108. *Csiszar I.* Information-type measures of difference of probability distributions and indirect observations // Stud. Sci. Math. Hungar. 1967. V. 2. P. 299–318.
109. *Csiszar I.* On the topological properties of  $f$ -divergence // ibidi. 1967. P. 329–339.
110. *Csiszar I.*  $I$ -divergence geometry of probability distributions and minimization problems // Ann. Probab. 1975. V. 3, № 1. P. 146–158.
111. *Darmonis G.* Sur les limites de la dispersion de certaines estimations // Rev. Inst. Intern. Statist. 1945. V. 13, № 1. P. 9–15.
112. *Dodson C.T.J.* (ed.) Geometrization of statistical theory // Proc. GST Workshop, Univ. of Lancaster Dept. of Mathem. — Lancaster: ULDM Publications, 1987.
113. *Dodson C.T.J.* Connection geometry // C.T.J. Dodson (ed.). Geometrization of Statistical Theory // Lancaster: Univ. Lanc. Dept. Math. Publ. 1987. P. 1–15.
114. *Dodson C.T.J., Jupp P.E., Kendall W.S., Lauritzen S.L.* A summary of points raised in the Discussion Sessions / ibidi // 1987. P. 235–250.
115. *Efron B.* Defining the curvature of a statistical problem (with applications to second order efficiency (with discussion) // Ann. Statist. 1975. V. 3. P. 1189–1242.
116. *Efron B.* The geometry of exponential families // Ann. Statist. 1978. V. 6, № 2. P. 362–376.

117. *Efron B.* Maximum likelihood and decision theory // *Ann. Statist.* 1982. V. 10, № 2. P. 340–356.
118. *Evans F., Fraser D.A.S., Monette G.* On principle and arguments to likelihood // *Can. J. Statist.* 1986. V. 14. P. 181–194.
119. *Fabian V., Hannan J.* On the Cramér–Rao inequality // *Ann. Statist.* 1977. V. 5. P. 197–205.
120. *Fisher R.A.* On the mathematical foundation of theoretical statistics // *Phil. Trans. Roy. Soc. London, Ser. A.* 1922. V. 222. P. 309–368.
121. *Fisher R.A.* Theory of statistical estimation // *Proc. Cambridge Phil. Soc.* 1925. V. 22. P. 700–725.
122. *Fisher R.A.* Two new properties of mathematical likelihood // *Proc. Roy. Soc. London, Ser. A.* 1934. V. 144. P. 285–307.
123. *Fisher R.A.* The logic of inductive inference // *J. Roy. Statist. Soc.* 1935. V. 98, № 1. P. 39–54.
124. *Fisher R.A.* Statistical methods and scientific inference. 3rd, ed. — N.Y.: Hafner, 1956.
125. *Foster B.L.* Differentiation on manifolds without a connection // *Mich. Math. J.* 1958. V. 5. P. 183–190.
126. *Foster B.L.* Higher derivatives in geometry and physics // *Proc. Roy. Soc. London, Ser. A.* 1989. V. 423. P. 443–455.
127. *Fraser D.A.S.* The Structure of Inference. — N.Y.: Wiley, 1968.
128. *Fraser D.A.S., Reid N.* On conditional inference for a real parameter: A differential approach on the sample space // *Biometrika.* 1988. V. 75. P. 251–264.
129. *Fréchet M.* Sur l’extension de certaines évaluations statistiques au cas de petits échantillons // *Rev. Inst. Intern. Statist.* 1943. V. 11. P. 183–205.
130. *Fu J.* Large sample point estimation: A large deviation theory approach // *Ann. Statist.* 1982. V. 10. P. 762–767.
131. *Fu J., Kass R.E.* A note on the interpretation of Bahedur bound and the rate of convergence of the maximum likelihood estimator // *Statist. Probab. Lett.* 1984. V. 2. P. 269–273.
132. *Ghash J.K., Subramaniam K.* Second-order efficiency of maximum likelihood estimators // *Sankhya. Ser. A.* 1974. V. 36. P. 325–358.
133. *Girshick M.A., Savage L.J.* Bayes and minimax estimates for quadratic loss functions // *Proc. 2nd Berkeley Symp. Math. Statist. Probab.* 1951. P. 53–73.
134. *Hajek J.* Asymptotically most powerfull rank-order tests // *Ann. Math. Statist.* 1962. V. 33. P. 1124–1147.
135. *Halmos P.R., Savage L.J.* Applications of the Radon–Nikodym theorem to the theory of sufficient statistics // *Ann. Math. Statist.* 1949. V. 20. P. 225–241.
136. *Hasminskii R.Z., Ibragimov I.A.* On asymptotic efficiency in the presence of an infinite-dimensional nuisance parameter // K. Ito and Yu.V. Prokhorov (eds). *Probab. Theory and Math. Statist.* // *Lect. Notes Math.* 1983. V. 1021. P. 195–229.
137. *Hinkley D.V.* Likelihood // *Can. J. Statist.* 1981. V. 8. P. 151–163.
138. *Hodges J.L., Lehmann E.L.* Some application of the Cramér–Rao inequality // *Proc. II Berkeley Symp. Math. Statist. Probab.* 1951. P. 13–22.
139. *Hodges J.L., Lehmann E.L.* Deficiency // *Ann. Math. Statist.* 1970. V. 41. P. 783–801.
140. *Hoeffding W.* The role of assumption in statistical decisions // *Proc. III Berkeley Symp. Math. Statist. Probab.* 1956. V. 1. P. 105–114.
141. *Hougaard P.* Parameterization of non-linear models // *J. Roy. Statist. Soc. Ser. B.* 1982. V. 44. P. 244–252.
142. *Ibragimov I.A., Has’minskii R.Z.* Statistical Estimation. Asymptotic Theory. — N.Y.: Springer, 1981.
143. *Ingarden R.S.* Information geometry in function spaces of classical and quantum finite statistical systems // *Int. Engrg. Sci.* 1981. V. 19. P. 1609–1633.
144. *Ingarden R.S., Sato Y., Sugawa K., Kawaguchi M.* Information thermodynamics and differential geometry // *Tensor, n. s.* 1979. V. 33. P. 347–353.

145. *Ireland C.T., Kullback S.* Minimum discrimination information estimation // *Biometrics*. 1968. V. 24, № 3.
146. *Jeffreys H.* An invariant form for the prior probability in estimation problems // *Proc. Roy. Soc. London. Ser. A*. 1946. V. 196. P. 453–461.
147. *Jeffreys H.* *Scientific inference*. 2nd ed. — Cambridge: Cambridge Univ. Press, 1957
148. *Jorgensen E.N.* Comparison of experiments when the parameter space is finite // *Z. Wahrscheinlichkeitstheor. und Verw. Geb.* 1970. V. 16. P. 219–249.
149. *Joshi V.M.* On attainment of the Cramér–Rao lower bound // *Ann. Statist.* 1976. V. 4. P. 998–1002.
150. *Jupp P.E.* Differential Geometry and parameters of interest / In: C.T.J. Dodson (ed.). *Geometrization of Statistical Theory* // Lancaster: Univ. Lanc. Dept. Math. Publ. 1987. P. 91–122.
151. *Kass R.E.* Canonical parametrizations and zero parameter-effects curvature // *J. Roy. Statist. Soc. Ser. B*. 1984. V. 46, № 1. P. 86–92.
152. *Kass R.E.* Introduction // *Differential geometry in statistical inference* // Hayward, Calif.: Inst. Math. Statist. 1987. P. 1–17.
153. *Kass R.E.* The geometry of asymptotic inference (with comments and rejoinder) // *Statist. Sci.* 1989. V. 4, № 3. P. 188–234.
154. *Kriz T.A., Talacko J.V.* Equivalence of the maximum likelihood estimator to a minimum entropy estimator // *Trab. Estadist. Invest. Oper.* 1968. V. 19, № 1–2. P. 55–65.
155. *Kullback S.* *Information Theory and Statistics*. — N.Y.: Wiley, 1959.
156. *Kullback S., Leibler R.A.* On information and sufficiency // *Ann. Math. Statist.* 1951. V. 22, № 1. P. 79–86.
157. *Kumon M., Amari S.-I.* Estimation of structural parameter in the presence of a large number of nuisance parameters // *Biometrika*. 1984. V. 71. P. 445–459.
158. *Kumon M., Amari S.-I.* Differential geometry of testing hypothesis: A higher order asymptotic theory in multiparameter curved exponential family // *J. Fac. Engrg., Univ. Tokyo, ser. B*. 1988. V. 39. P. 241–273.
159. *Lauritzen S.L.* *Statistical manifolds* // *Differential Geometry in Statistical Inference* // Hayward, Calif.: Inst. Math. Statist. 1987. P. 163–216.
160. *Lauritzen S.L.* Conjugate connections in statistical theory / C.T.J. Dodson (ed.). *Geometrization of Statistical Theory*. — Lancaster: Univ. Lanc. Dept. Math. Publications, 1987. — P. 33–51.
161. *Le Cam L.* On the asymptotic theory of estimation and testing hypothesis // *Proc. 3rd Berkeley Symp. Math. Statist. Probab.* 1956. V. 1. P. 129–156.
162. *Le Cam L.* Sufficiency and approximate sufficiency // *Ann. Math. Statist.* 1964. V. 35, № 4. P. 1419–1455.
163. *Le Cam L.* Likelihood functions for large numbers of independent observations // F.N. David (ed.). *Research Papers in Statistics. Festschrift for J. Neyman*. — London: Wiley, 1966.
164. *Le Cam L.* *Asymptotic methods in statistical decision theory*. — N.Y.: Springer, 1986.
165. *Lehmann E.L.* *Theory of Point Estimation*. — N.Y.: Wiley, 1983.
166. *McCullagh P., Cox D.R.* Invariance and likelihood ratio statistics // *Ann. Statist.* 1986. V. 14. P. 1419–1430.
167. *Mora M.* Geometrical expansions for the distributions of the score vector and the maximum likelihood estimator / Research Report № 172 Dept. Theor. Statist. — Aarhus Univ., 1988
168. *Morozova E.A., Čencov N.N.* Markov maps in noncommutative probability theory and mathematical statistics // Yu.V. Prohorov et al. (eds). *Probab. Theor. and Math. Statist. Proc. 4th Vilnius Conf. Utrecht* // VNU Sci. Press. 1987. V. 2. P. 287–310.
169. *Morse N., Sacksteder R.* Statistical isomorphism // *Ann. Math. Statist.* 1966. V. 37, № 1. P. 203–214.
170. *Neveu J.* *Bases mathématiques du calcul des probabilités*. — Paris: Dunod, 1964.
171. *Neyman J., Pearson E. S.* On the use and interpretation of certain test criteria for purposes of statistical inference // *Biometrika*. 1928. V. 20A. P. 175–240.

172. *Onicescu O.* Energie informationelle // C. r. Acad. sci. Paris. 1966. V. AB263, № 22. P. 841–842.
173. *Parzen E.* On estimation of a probability density function and mode // Ann. Math. Statist. 1962. V. 33, № 3. P. 1065–1076.
174. *Pfanzagl J.* First order efficiency implies second order efficiency // J. Jureckova (ed.). Contributions to Statistics J. Hajek Memorial Volume. — Prague: Academia, 1979. — P. 167–196.
175. *Pfanzagl J.* Contribution to a general asymptotic statistical theory // Lect. Notes Statist. — N.Y.: Springer. 1982. V. 13.
176. *Picard D.B.* Invariance properties of Fisher–Rao metric and Chentsov–Amari connections in regular families. — Republ. Univ. Paris-Sud, 1987.
177. *Picard D.B.* Invariance properties of metrics and connections in regular families // C.T.J. Dodson (ed.). Geometrization of Statistical Theory. // Lancaster: Univ. Lanc. Dept. Math. Publ., 1987. — P. 203–208.
178. *Pitman E.J.G.* Some basic theory for statistical inference. — London: Chapman and Hall, 1979.
179. *Rao C.R.* Information and accuracy attainable in the estimation of statistical parameters // Bull. Calcutta Math. Soc. 1945. V. 37. P. 81–91.
180. *Rao C.R.* Efficient estimates and optimum inference procedures in large samples (with discussion) // J. Roy. Statist. Soc. Ser. B. 1962. V. 24, № 1. P. 46–72.
181. *Rao C.R.* Differential metrics in probability spaces // Differential geometry in statistical inference. — Hayward, Calif.: Inst. Math. Statist., 1987.
182. *Rao C.R.* Comment on R.E. Kass the geometry of asymptotic inference // Statist. Sci. 1989. V. 4, № 3. P. 229–231.
183. *Rao C.R., Sinha B.K., Subramanyam K.* Third order efficiency of the maximum likelihood estimator in the multinomial distribution // Statist. and Decisions. 1982. V. 1. P. 1–16.
184. *Reid N., Fraser D.A.S.* Comment on R. E. Kass The Geometry of Asymptotic Inference // Statist. Sci. 1989. V. 4, № 3. P. 231–233.
185. *Rényi A.* On measures of entropy and information // Proc. 4th Berkeley Symp. Math. Statist. Probab. // 1961. V. 1. P. 537.
186. *Rockafellar R.T.* Characterization of the subdifferentials of convex functions // Pacif. J. Math. 1966. V. 17, № 3. P. 497–510.
187. *Romier G.* Exhaustivité et equivalence des objets statistiques // C. r. Acad. sci. Paris. 1968. V. 267, № 22. P. 828–831.
188. *Sacksteder R.* A note on statistical equivalence // Ann. Math. Statist. 1967. V. 38, № 3. P. 787–795.
189. *Sacksteder R.* On products of experiments // Z. Wahrscheinlichkeitstheor. und verw. Geb. 1968. V. 10, № 3. P. 203–211.
190. *Schwartz S.C.* Estimation of probability density by an orthogonal series // Ann. Math. Statist. 1967. V. 38, № 4. P. 1261–1265.
191. *Sen R.N.* On parallelism in Riemannian space. I, II, III // Bull. Calcutta // Math. Soc. 1944. V. 36. P. 102–107 // 1945. V. 37. P. 153–159 // 1946. V. 38. P. 161–167.
192. *Skovgaard L.T.* A Riemannian geometry of the multivariate normal model // Scand. J. Statist. 1984. V. 11. P. 211–223.
193. *Stein Ch.* Notes on comparison of experiments / Mimeographed. — Univ. Chicago, 1951.
194. *Van Ryzin J.* Bayes risk consistency of classification procedure using density estimation // Sankhya, ser. A. 1966. V. 28. P. 261–270.
195. *Wald A.* Statistical decision function. — N.Y.: Wiley, 1950.
196. *Watson G.S.* Density estimation by orthogonal series // Ann. Math. Statist. 1969. V. 40, № 4. P. 1496–1498.
197. *Wijsman R.A.* On the attainment of the Cramer–Rao lower bound // Ann. Statist. 1973. V. 1. P. 538–542.
198. *Witting H.* Mathematische statistik, I. — Stuttgart: Teubner, 1985.

## ПРОЕКТИВНО-ЕВКЛИДОВА ГЕОМЕТРИЯ И НЕКОММУТАТИВНАЯ ТЕОРИЯ ВЕРОЯТНОСТЕЙ<sup>\*)</sup>

Замечательным свойством Бориса Николаевича Делоне было его умение смотреть на задачи алгебры и теории чисел глазами геометра, находя к ним новые неожиданные подходы. Это качество он всегда пытался привить и своим ученикам. В настоящей работе мы предпринимаем попытку чисто геометрического анализа понятия квантовых логик — решеток линейных подпространств евклидова или гильбертова пространства, замкнутых относительно некоторых геометрических операций. Показывается, что зеркальные симметрии относительно подпространств квантовой логики порождают клиффордово исчисление гомотетий, которое можно использовать для координатизации секторных связей атомов решетки. Для введения умножения гомотетий использованы конфигурация Паша для сферических треугольников и конфигурация теоремы о трех перпендикулярах. Развитые конструкции позволяют согласовать решеточный подход Биркгофа-фон Неймана [1] и алгебраический подход Иордана-фон Неймана-Вигнера [2], чего ранее не удавалось [3].

1. Случайные явления микромира не описываются схемами классической (коммутативной, как теперь говорят) теории вероятностей, потому что логика квантовых событий не является булевой [4, 5]. Описание же такого явления, как объекта некоммутативной теории вероятностей, удобно давать следующим образом (ср. [5]). задается сепарабельное абстрактное гильбертово пространство  $X$  объекта. Его единичные векторы  $x \in X$  или по крайней мере некоторые из них описывают элементарные исходы явления (чистые состояния объекта). задается инъективная ультраслабо замкнутая йорданова алгебра  $A$  ограниченных эрмитовых операторов  $A$ , действующих на  $X$ , — алгебра ограниченных наблюдаемых (аналогов случайных величин). Вероятностное состояние объекта описывается неотрицательным нормированным нормальным  $\mathbb{R}$ -линейным функционалом  $\Phi$  на  $A$ . Этот функционал в рамках описываемой схемы удобно задавать ядерным эрмитовым оператором  $\Phi \geq 0$  на  $X$ ,  $\text{tr } \Phi = 1$ . Тогда среднее значение, т.е. математическое ожидание наблюдаемой  $A$ , при состоянии  $\Phi$  описывается величиной  $\text{tr } A\Phi = \text{tr } \Phi A$ . Событиям (“да-нет” экспериментам) отвечают идемпотенты алгебры  $A$  — ортопроекторы на подпространства  $X$ . Элементарные исходы явления, как видим, не являются событиями, равно как и чистые (векторные) состояния не являются вероятностными состояниями схемы.

Если йорданова алгебра  $A$  есть самосопряженная часть своей обертывающей алгебры фон Неймана  $B$ , мы будем называть систему всех идемпотентов из  $A$  *неймановой логикой*. В противном случае мы будем называть такую систему *йордановой логикой*.

2. Булевы алгебры событий теории вероятностей принято определять конструктивно как систему, замкнутую относительно операций объединения, пе-

<sup>\*)</sup> Труды МИАН. 1991. Т. 196. С. 105–113 (совм. с Е.А. Морозовой).

ресечения и собственного вычитания (объемлемого события из объемлющего), а также операции счетного объединения (или перехода к монотонному пределу). Для конечных алгебр последняя операция излишня. Когда алгебра унитарна, т.е. обладает максимальным элементом, то бинарную операцию вычитания можно заменить на унарную операцию перехода к дополнению. Соответственно этой схеме Биркгоф и фон Нейман определили [1] квантовые логики событий операционно как системы (ультраслабо замкнутых) подпространств, замкнутые относительно операций линейной суммы двух подпространств  $(L, K) \rightarrow L \vee K$ , пересечения двух подпространств  $(L, K) \rightarrow L \cap K$ , перехода к ортогональному дополнению  $L \rightarrow L' = X \ominus L$  и перехода к монотонному пределу. Класс таких систем — недистрибутивных ортодополнительных решеток — оказывается слишком широким. Чтобы выделить из него класс йордановых логик, надо добавить к указанным Биркгофом и фон Нейманом базовым операциям еще две операции, не имевшие прямых классических аналогов.

Обратимся еще раз к классической схеме. Естественно назвать перестановочным ковариантом некоторого класса  $\mathcal{C}$  подмножеств  $L_i$ ,  $i = 1, 2, \dots$ , конечного множества  $E$  всякое множество  $K$ , которое переходит в себя,  $T(K) = K$ , при любой перестановке  $T$  элементов тотального множества  $E$ , переводящей в себя каждое из подмножеств  $L_i$ ,  $T(L_i) = L_i$ ,  $i = 1, 2, \dots$ . Очевидно, что операции объединения, пересечения и разности двух множеств являются ковариантными, сопоставляя паре некий ее ковариант. Нетрудно проверить, что эта система операций полна, т.е. любой перестановочный ковариант для класса  $\mathcal{C}$  можно построить с помощью суперпозиций бинарных операций, отправляясь от исходных  $\mathcal{C}$ -подмножеств, например, по дизъюнктивной нормальной форме.

**3.** Оказывается, для конечномерных гильбертовых пространств можно поступить сходным образом. Назовем *унитарным* (соответственно *унитарно-антиунитарным*) *ковариантом* некоторого класса  $\mathcal{C}$  линейных подпространств  $L_1$  конечномерного гильбертова пространства  $X$  всякое подпространство  $K$ , которое переходит в себя,  $U(K) = K$ , при любом унитарном (соответственно унитарном или антиунитарном) преобразовании  $U$  пространства  $X$ , переводящем в себя каждое из подпространств  $L_i$ ,  $U(L_i) = L_i$ ,  $i = 1, 2, \dots$ .

Условимся далее для упрощения обозначать ортопроектор на подпространство  $L$  той же буквой; тогда ортопроектор на  $U(L)$  будет  $ULU^{-1}$ .

**ТЕОРЕМА 1.** *Совокупность всех унитарных ковариантов для класса  $\{L_1, L_2, \dots\}$  будет неймановой логикой, отвечающей минимальной алгебре фон Неймана, содержащей ортопроекторы  $L_1, L_2, \dots$*

**ТЕОРЕМА 2.** *Совокупность всех унитарно-антиунитарных ковариантов для пары  $\{L_1, L_2, \dots\}$  будет йордановой логикой, отвечающей минимальной йордановой алгебре, содержащей ортопроекторы  $L_1$  и  $L_2$ .*

Для того чтобы унитарное преобразование  $U$  переводило подпространство в себя, необходимо и достаточно, чтобы  $ULU^{-1} = L$ , т.е. чтобы  $U$  и  $L$  коммутировали. Поэтому утверждение теоремы 1 следует из теоремы фон Неймана о бикоммутанте. При этом неймановы логики порождаются уже бинарными унитарно ковариантными операциями. Антиунитарно ковариантные из них порождают йордановы логики (см. п. 5, 7, 9). Полная система унитарно-антиунитарно

ковариантных операций включает еще 4-арную операцию и порождает вигнеровы логики, отвечающие реверсным йордановым алгебрам.

4. Изучим теперь бинарные ковариантные операции в квантовых логиках.

ОПРЕДЕЛЕНИЕ 1. Подпространства  $F$  и  $G$  называются *изоклинными* друг другу (точнее, наклоненными друг к другу под постоянным углом  $\varphi$ ), если для их ортопроекторов

$$FGF = \cos^2 \varphi \cdot F, \quad GFG = \cos^2 \varphi \cdot G, \quad \cos^2 \varphi \neq 0, \quad (1)$$

т.е. каждый вектор одного из подпространств образует со своей ортопроекцией на второе подпространство указанный угол  $\varphi$ . Этот факт мы будем записывать  $(F, G) = \cos^2 \varphi$  [6, 7].

В общем случае для пары подпространств  $L$  и  $K$  существует ортогональное разложение Жордана [8]

$$L = (L \cap K) \oplus L_1 \oplus \dots \oplus L_k \oplus L_{k+1}, \quad (2)$$

$$K = (L \cap K) \oplus K_1 \oplus \dots \oplus K_k \oplus K_{k\pm 1}, \quad (3)$$

$$L \vee K = (L \cap K) \oplus (L_1 \cap K_1) \oplus \dots \oplus (L_k \cap K_k) \oplus (L_{k+1} \cap K_{k+1}), \quad (4)$$

где все пары  $L_j, K_j$  изоклинные,  $L_{k+1} = L \cap K'$ ,  $K_{k+1} = K \cap L'$ ,  $1 > (L_1, K_1) > \dots > (L_k, K_k) > 0$ ,  $L'$  и  $K'$  — ортогональные дополнения к  $L$  и  $K$ . Очевидно, все элементы этого разложения являются унитарно-антиунитарными (тем более унитарными) ковариантами пары подпространств  $L$  и  $K$ . Ясно также, что средние члены  $L$ - и  $K$ -разложения (если они отличны от нулевого пространства) операциями Биркгофа и фон Неймана не получить. Аккуратное формальное доказательство их неполноты, проведенное на еще более простом примере, см. в работе [9, формула (0.10)].

Нетрудно видеть, что величины  $(L_j, K_j)$  являются общими собственными числами  $\lambda_j$  операторов  $LKL$  и  $KLK$ . Значение  $\lambda_0 = 1$  принадлежит спектру тогда и только тогда, когда  $L \cap K \neq \{0\}$ . Имеющую смысл синуса угла величину  $(1 - \lambda_{\max})^{1/2}$  мы будем называть *щелью*  $\omega(L, K)$  между  $L$  и  $K$ , а  $(1 - \lambda_{\min})^{1/2}$  — *раствором*  $d(L, K)$ . Таким образом, между пересекающимися подпространствами и только между ними щель равна нулю.

ОПРЕДЕЛЕНИЕ 2. Назовем *контактом* подпространства  $L$  с неортогональным ему подпространством  $K$  (т.е. при  $\omega(L, K) < 1$ ) собственное подпространство  $L| \cap K \subseteq L$  эрмитова оператора  $LKL$ , отвечающее  $\lambda_{\max}$ .

Таким образом, контакт  $L| \cap K$  совпадает с  $L \cap K$  и с контактом  $K| \cap L$ , когда  $L \cap K \neq \{0\}$ . В противном случае эти контакты  $L| \cap K$  и  $L \cap |K$  изоклинные. Контакты ортогональных подпространств считаются нулевым подпространством по определению. Так что всегда ортопроектор  $P$  на контакт  $L| \cap K$  будет

$$P = \lim_{n \rightarrow \infty} (L, K)^{-n} (LKL)^n, \quad (5)$$

что обобщает выражение для ортопроектора на пересечение.

5. Проектирование одного изоклинного подпространства на другое задает соответствующую им гомотегию с коэффициентом сжатия  $\cos \varphi$ . Эти гомотегии определяют изометрию  $I$  подпространства  $F$  на  $G$  и обратно,  $I(G \leftarrow F)$ ,

$I(F \leftarrow G)$ :

$$x \rightarrow Ix = y = (\cos \varphi)^{-1}Gx, \quad y \rightarrow x = (\cos \varphi)^{-1}Fy, \quad (6)$$

последняя может быть задана зеркальной симметрией  $M$  всего пространства

$$M = 2H - 1, \quad M|_F = I(G \leftarrow F), \quad M|_G = I(F \leftarrow G), \quad (7)$$

относительно их биссектора  $H = \{z = x + Ix \ \forall x \in F\}$  с ортопроектором

$$H = (4 \cos^2 \varphi)^{-1}[F + G - (2 \cos^2 \varphi - 1)^{-1}(FG + GF)] \quad (8)$$

или же частичной симметрией

$$N = 2H - B, \quad (9)$$

где  $B$  — ортопроектор на  $F \vee G$ .

**ОПРЕДЕЛЕНИЕ 3.** Назовем  $(a : b)$  *суперпозицией* (или когерентной комбинацией) изоклинных подпространств  $F$  и  $G$  векторное подпространство

$$S(aF : bG) = \{z : z = ax + bNx \ \forall x \in F\}. \quad (10)$$

Мы будем различать  $\mathbb{R}$ -суперпозиции и  $\mathbb{C}$ -суперпозиции в зависимости от числовой природы отношения  $a : b$ .

Очевидно, все  $\mathbb{R}$ -суперпозиции (соответственно все  $\mathbb{C}$ -суперпозиции) изоклинной пары  $F, G$  образуют пучок, изоморфный (или физически эквивалентный в смысле [5]) пучку прямых вещественной (соответственно комплексной) плоскости, в котором пара  $a : b$  задает однородные координаты. Очевидно,  $H$  является  $(1 : 1)$ -суперпозицией. Ортопроекторы на  $S(aF : bG)$  выражаются линейным образом через  $F, G$  и  $H$  (или  $FG + GF$ ). Наконец, суперпозиции двух любых неортогональных подпространств пучка дают снова тот же пучок.

**ОПРЕДЕЛЕНИЕ 4.** Назовем класс  $\Lambda$  подпространств  $\mathbb{R}$ -логикой (соответственно  $\mathbb{C}$ -логикой), если он замкнут относительно бинарных операций линейной суммы  $L \vee K$ , контакта  $L \cap K$ , ортогонального вычитания  $L \ominus K$  для пар  $L \supseteq K$  и  $\mathbb{R}$ -суперпозиции (соответственно  $\mathbb{C}$ -суперпозиции) (когерентного комбинирования)  $S(aF : bG)$  для изоклинных пар.

Можно спорить, насколько удачно выбрана система операций. Особенно неожиданно, что операция когерентного комбинирования определена только для изоклинных пар. Однако с похожей ситуацией приходится сталкиваться, как мы впервые узнали из университетских лекций Бориса Николаевича Делоне, например, в теории кривых второго порядка, вводя семиинварианты (см. [10, § 108]). Конечно, в бесконечномерном случае надо заменять и операцию контакта, заменяя ее на некие спектральные интегралы от проекторов по общему спектру  $LKL$  и  $KLK$ .

Нетрудно показать, что любой унитарно-антиунитарный (соответственно унитарный) ковариант  $Q$  пары  $L$  и  $K$  общего положения также разложим по Жордану:

$$Q = Q_0 \oplus Q_1 \oplus \dots \oplus Q_k \oplus Q_{k+1} \oplus Q_{k+2}, \quad (11)$$

где  $Q_0 = L \cap K$  или  $Q_0 = \{Q\}$ ,  $Q_j = L_j \vee K_j$ , или  $Q_j = S(aL_j : bK_j)$  с некоторым

отношением  $(a : b) \in \mathbb{R}$  (соответственно  $(a : b) \in \mathbb{C}$ ), или  $= \{0\}$ ,  $Q_{k+1} = (L \cap \cap K') \oplus (L' \cap K)$ , или  $Q_{k+1} = L \cap K'$ , или  $Q_{k+1} = L' \cap K$ , или  $Q_{k+1} = \{0\}$ ,  $Q_{k+2} = L' \cap K'$ , или  $Q_{k+2} = \{0\}$  (см. [11, теорема 7.2]). Очевидно, перечисленных в определении 4 операций достаточно, чтобы построить любой из них исходя из  $L$  и  $K$  (и еще  $X$  для  $L' \cap K'$ ). Другими словами, для построения бинарных ковариантов наших операций хватает. С другой стороны, все указанные  $\mathbb{R}$ -коварианты принадлежат йордановой логике, порожденной проекторами  $L$  и  $K$ . Это следует (см. [11, теорема 11.1]) из разложения Пирса [2, 5].

**6.** В конечных булевых логиках очень удобно использовать дизъюнктивную нормальную форму, представляющую и притом единственным образом элемент логики в виде объединения атомов, т.е. минимальных элементов логики. Подобные представления, но уже неединственные, существуют и в квантовых логиках.

**ОПРЕДЕЛЕНИЕ 5.** Элемент  $A$  логики  $\Lambda$  называется *минимальным подпространством* или *атомом логики*, если  $A \supseteq B \in \Lambda$  влечет: или  $B = \{0\}$ , или  $B = A$ .

Из определений 5 и 4 вытекает, что любой элемент конечномерной логики разложим в ортосумму атомов.

**ЛЕММА 1.** *Ортопроекция атома на любое подпространство логики или  $\{0\}$ , или снова атом. Любые два атома или ортогональны, или изоклинны.*

**ОПРЕДЕЛЕНИЕ 6.** Назовем два атома логики *связуемыми*, если они контактируют (и тем самым изоклинны) или ортогональны, но существует третий атом, контактирующий с каждым из них.

**ЛЕММА 2.** *Отношение связуемости рефлексивно, симметрично и транзитивно.*

Доказательство леммы вытекает из теорем 1.4 и 1.5 работы [11]. Таким образом, возникает разложение

$$X = Y_1 \oplus \dots \oplus Y_m \oplus X_0 \tag{12}$$

исходного пространства в сумму “секторов”  $Y_1, \dots, Y_m$ , где  $Y_1 \oplus \dots \oplus Y_m$  — максимальный элемент логики, и  $X_0 = \{0\}$ , когда само  $X$  — элемент логики. При этом аналогичное разложение выполнено для любых подпространств  $L$  логики, так что для ортопроектора

$$L = Y_1 L Y_1 \oplus \dots \oplus Y_m L Y_m. \tag{13}$$

Может оказаться, что некоторые из секторов  $Y_j$  являются атомами. Тогда их объединение порождает логику, изоморфную булевой. Поэтому интересен только случай неминимальных секторов.

**7.** Рассмотрим теперь предминимальное подпространство логики, в котором по определению ровно два ортогональных атома, скажем,  $F$  и  $G$ . Проведем  $\mathbb{R}$ -пучок через  $F$  и любой третий атом  $H$ . Он целиком лежит в  $F \vee H = F \oplus G$ . При некотором значении  $a : b$ , как нетрудно проверить, формула (10) задает пространство, ортогональное  $F$ , т.е.  $G$ . Не ограничивая общности, можно считать, что  $(F, H) = 1/2$ , т.е. угол между  $F$  и  $H$  равен  $\pi/4$ . Тогда  $(G, H) = 1/2$ , и общий вид ортопроектора на элемент пучка, образующий с  $F$  угол  $\varphi$ , есть

$$P(\varphi) = \cos^2 \varphi \cdot F + \sin \varphi \cdot \cos \varphi \cdot NF + \sin \varphi \cdot \cos \varphi \cdot NF + \sin^2 \varphi \cdot G \tag{14}$$

для  $\mathbb{R}$ -пучка, где  $0 \leq \varphi \leq \pi$ ,

$$N = 2H - (F \oplus G), \quad (15)$$

угол  $\varphi$  отсчитывается так, что  $P(\pi/4) = H$ , и

$$P(\varphi, \theta) = \cos^2 \varphi \cdot F + \sin \varphi \cdot \cos \varphi \cdot e^{i\theta} NF + \sin \varphi \cdot \cos \varphi \cdot e^{-i\theta} NG + \sin^2 \varphi \cdot G \quad (16)$$

для  $\mathbb{C}$ -пучка,  $0 \leq \theta \leq \pi$ .

“Срединный” атом  $H$  и отвечающая ему частичная зеркальная изометрия  $N$  устанавливают изометрию  $F$  и  $G$ . К той же изометрии мы придем, если будем проектировать  $F$  на  $H$ , а затем  $H$  на  $G$  (и удваивать результат):

$$\forall x \in F \quad Nx = 2GHx, \quad x = 2FH(Nx), \quad (17)$$

или использовать в качестве промежуточного любой атом  $P(\varphi)$  пучка при  $0 < \varphi < \pi/2$ . Вторая половина  $\mathbb{R}$ -пучка вместе с атомом  $P(3\pi/2) = (F \oplus G) \ominus H$  дает изометрию  $-N$ .

**ЛЕММА 3.** *Через два ортогональных связуемых атома  $\mathbb{C}$ -логики проходит только один  $\mathbb{C}$ -пучок атомов. Все атомы, лежащие в одном предминимальном пространстве, образуют один  $\mathbb{C}$ -пучок.*

**Доказательство.** Пусть через  $F$  и  $G$  проходят два  $\mathbb{C}$ -пучка с зеркальными изометриями  $N_1$  и  $N_2$ . Тогда  $V = N_1 N_2$  задает унитарное преобразование  $F$  на себя. Пусть  $x \in F$  есть собственный вектор,  $Vx = e^{i\theta} x$ . Тогда срединные атомы  $H_1$  и  $H_2$  пересекаются по вектору  $x + e^{i\theta} N_1 x = x + N_1 Vx = x + N_2 x$ . Следовательно, атомы совпадают  $V = e^{i\theta}$  всюду на  $F$ ,  $e^{i\theta} I_1 = I_2$ .

**ЛЕММА 4.** *Пусть атом  $G$  контактирует с атомом  $E \subset D \in \Lambda$ . Тогда ортопроекция  $K$  атома  $G$  на подпространство  $D$  будет атомом, изоклинным как с  $E$ , так и с  $G$ , причем изометрия (6) атома  $G$  на  $E$  совпадает с произведением изометрий  $G$  на  $K$  и  $K$  на  $E$ :*

$$I(E \leftarrow G) = I(E \leftarrow K) I(K \leftarrow G), \quad (18)$$

$$(E, G) = (E, K) (K, G). \quad (19)$$

**СЛЕДСТВИЕ.** *Если в условиях леммы 4  $D \supseteq (E \vee F)$ , где  $F$  — атом, также контактирующий с  $E$ , то  $K$  изоклинно и с  $F$ , причем*

$$I(F \leftarrow G) I(G \leftarrow E) = I(F \leftarrow K) I(K \leftarrow E), \quad (20)$$

$$(E, G) : (F, G) = (E, K) : (F, K). \quad (21)$$

**Доказательство.** Для всякого  $y \in G$  по (6)

$$\begin{aligned} I(E \leftarrow G)y &= (E, G)^{-1/2} E y = (E, G)^{-1/2} E K y = \\ &= (EG)^{-1/2} (EK)^{-1/2} I(E \leftarrow K) I(K \leftarrow G) y. \end{aligned}$$

Так как изометрия не изменяет длины, то последний множитель равен 1. Так как при обращении стрелок возникают обратные изометрии, то это дает (20) из (18).

Использованное равенство  $EG = EKG$  при  $K = DG$  есть следствие евклидовой теоремы о трех перпендикулярах: “проекция на  $E$  есть проекция на  $E$  проекции на объемлющее  $D$ ”.

Теперь мы можем ввести в сектор  $Y \wedge Y$   $\mathbb{C}$ -логики координатизацию. Пусть  $Y = E_1 \oplus \dots \oplus E_k$  — какое-то разложение сектора  $Y$  на ортогональные атомы,  $F$  — атом, контактирующий со всеми  $E_j$ .

**ТЕОРЕМА 3.** *Совокупность всех атомов  $\mathbb{C}$ -логики  $\Lambda$  в секторе  $Y$  изоморфна связке всех прямых  $k$ -мерного комплексного гильбертова пространства. Определемому (с точностью до множителя  $s \neq 0$ ) набором однородных координат  $z_1 : z_2 : \dots : z_k$  относительно репера  $E_1, \dots, E_k, F$  отвечает ровно один атом, и обратно.*

Для доказательства достаточно заметить, что за однородные координаты атома  $G$  можно в силу следствия из леммы 4 взять набор  $z_1 : z_2 : \dots : z_k$ , где все  $|z_j| = (G, E_j)^{1/2}$ ,  $\arg z_j = \theta_j$ , а проекция  $G_{0j}$  атома  $G$  на  $E_j \vee F = B_a \theta_j$  задается по правилу:  $\theta_j = \text{Im} \ln \lambda_j$ , где  $\lambda E_j = E_j F (B \ominus E_j) G_{0j} E_j$ .

Из теоремы 3 вытекает независимое доказательство теоремы 1.

**8.** Структура системы атомов предминимального пространства  $\mathbb{R}$ -логики может быть более сложной. Ее естественное описание будет дано в терминах векторного представления спинорной группы [13].

Пусть ортогональное разложение предминимального пространства  $B = E \oplus \oplus F$  фиксировано. Каждый атом  $G$  в  $B$  входит в один и только один  $\mathbb{R}$ -пучок, проходящий через  $E$  и  $F$ . Каждый такой пучок может быть охарактеризован одним из двух своих срединных атомов  $H$  и  $B \ominus H$ , образующих с  $E$  и  $F$  углы  $\pi/4$  (см. (14) и (15)). Точно так же, как для изоклинных  $E$  и  $F$ , можно рассматривать  $H$  как биссектор (один из многих) прямого угла между  $E$  и  $F$  и связывать с  $H$  не только зеркальную симметрию  $M_H$  всего пространства  $X$ , но и зеркальную симметрию  $N_H = 2H - B$  подпространства  $B$ .

**ЛЕММА 5.** *Пусть  $K$  и  $H$  — срединные атомы. Тогда*

$$N_H N_K + N_K N_H = [2(H, K) - 1] B = [N_H, N_K] B, \quad (22)$$

где  $2(H, K) - 1 = \cos 2\varphi$ ,  $\varphi$  — угол между  $K$  и  $H$ .

**ЛЕММА 6.** *Пусть  $K$  и  $H$  не совпадают и неортогональны. Тогда любая  $(a : b)$  — суперпозиция атомов  $K$  и  $Y$ ,  $(a : b) \in \mathbb{R}$ , — также будет срединным атомом  $G$  с зеркальной симметрией*

$$N_G = \alpha N_K + \beta N_H, \quad \alpha, \beta \in \mathbb{R}. \quad (23)$$

*Обратно, каждому значению отношения  $\alpha : \beta$  отвечают две суперпозиции  $K$  и  $H$  с нормировкой*

$$1 = \alpha^2 + 2\alpha\beta[N_K, N_H] + \beta^2. \quad (24)$$

Доказательство этих лемм проводится с использованием блочных матриц, описывающих операторы в  $B = E \oplus F$  (см. [11], где в формулировке леммы 12.2 перепутано, что однозначно. Все выкладки в доказательстве там сделаны правильно).

Фиксируем теперь какую-нибудь зеркальную симметрию  $N_0$  как “начальную” и рассмотрим все индуцированные симметрии вида  $N_0 N_H$  атома  $E$  на себя, а также гомотетии  $\lambda N_0 N_H = W$ .

ТЕОРЕМА 4. Зеркальные симметрии, порождаемые  $\mathbb{R}$ -логикой в предминимальном пространстве с фиксированным разложением  $B = E \oplus F$ , порождают векторное пространство  $W$  гомотетий атома  $E$  на себя вида  $W = \lambda N_0 N_H$  с евклидовой структурой

$$(\lambda N_0 N_H, \mu N_0 N_K) = \lambda \mu [N_H, N_K] \quad (25)$$

и структурой йордановой алгебры

$$WZ + ZW = [W, B]Z + [Z, B]W - [W, Z]B. \quad (26)$$

Все элементы единичной длины обратимы в  $B$ :

$$(N_0 N_H)^{-1} = N_H N_0 = [N_H, N_0]B - N_0 N_H. \quad (27)$$

СЛЕДСТВИЕ. Указанные гомотетии образуют алгебру Клиффорда.

Доказательство вытекает из утверждений лемм 5 и 6. Таким образом, конструкции, проведенные в этом пункте, позволяют задавать атомы  $A$  предминимального пространства клиффордовыми числами  $\gamma_0 e_0 + \dots + \gamma_n e_n = \Gamma$ , где  $[\gamma_0^2 + \dots + \gamma_n^2]^{1/2} = \operatorname{tg} \varphi$ ,  $\varphi$  — угол наклона  $A$  к  $E$ , единичный вектор  $(\operatorname{tg} \varphi)^{-1} \Gamma$  задает положение срединного атома пучка, проходящего через  $E$  и  $A$ , по отношению к выбранным “ортогональным” в смысле (22) базисным срединным атомам  $H_0, H_1, \dots, H_n$ . Тем самым мы дали полное описание всех предминимальных секторов.

9. Мы видели, что зеркально порожденные гомотетии можно складывать. А вот последовательное выполнение пар зеркальных симметрии уже, вообще говоря, выводит из указанного векторного пространства. Оказывается, что в секторных подпространствах формальной размерности три и выше (т.е. не являющихся ни атомами, ни предминимальными подпространствами) такое умножение сохраняет геометрический тип.

Этот факт устанавливается с помощью ряда лемм. Все они фактически используют, что большой круг, пересекающий две стороны сферического треугольника и не проходящий через вершины, пересекает продолжение третьей стороны. Всюду далее  $D = E_1 \oplus E_2 \oplus E_3$  лежит в одном секторе логики  $\Lambda$ .

ЛЕММА 7. Клиффордовы алгебры гомотетий атома  $E_1$ , индуцированные подлогиками  $\Lambda \cap (E_1 \oplus E_2)$ ,  $\Lambda \cap (E_1 \oplus E_3)$ , геометрически изоморфны.

Доказательство вытекает из леммы 4 и следствия из нее.

ЛЕММА 8. Пусть  $G_{12} \subset (E_1 \oplus E_2)$ ,  $G_{12} \neq E_1, E_2$ ,  $G_{23} \subset (E_2 \oplus E_3)$ ,  $G_{23} \neq E_2, E_3$ ,  $G_{12}$  и  $G_{23}$  контактируют. Тогда проходящий через них  $\mathbb{R}$ -пучок пересекает  $E_3 \oplus E_1$  по атому  $G_{31} \neq E_1, E_3$ ,  $G_{31} = (G_{12} \vee G_{23}) \cap (E_1 \oplus E_3)$ . Все изометрии  $J(E_j \leftarrow E_i)$ , определяемые через  $G_{ij}$ , связаны соотношениями

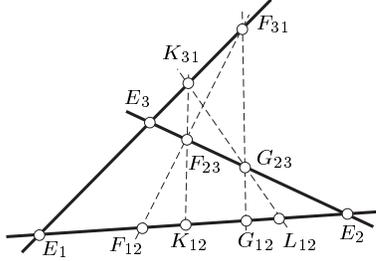
$$J(E_3 \leftarrow E_1) = -J(E_3 \leftarrow E_2) J(E_2 \leftarrow E_1), \quad (28)$$

$$J(E_j \leftarrow E_i) = I(E_j \leftarrow G_{ij}) I(G_{ij} \leftarrow E_i). \quad (29)$$

ЛЕММА 9. Пусть  $H_{12} \subset (E_1 \oplus E_2)$ ,  $H_{23} \subset (E_2 \oplus E_3)$  определяют “начальные” зеркальные отображения  $E_2$  на  $E_1$  и  $E_2$  на  $E_3$ , и пусть  $W$  и  $Z$  — изометрии пространства  $E_1$  на себя, определяемые атомами  $G_{12}$ ,  $K_{12} \subset (E_1 \oplus E_2)$  по схеме

теоремы 4. Тогда найдется атом  $L_{12} \subset (E_1 \oplus E_2)$ , определяющий таким же образом изометрию  $WZ$ .

Доказательство получается (см. [11]) применением формул (28) и (29)



при следующем построении. Проводим через  $F_{12}$  и  $F_{23}$  пучок, пересекающий  $E_1 \oplus E_3$  по атому  $F_{31}$ . Проведем через  $G_{12}$  и  $F_{31}$  новый пучок, пересекающий  $E_2 \oplus E_3$  по атому  $G_{23}$ , а также через  $F_{23}$  и  $K_{12}$  пучок, пересекающийся с  $E_1 \oplus E_3$  по атому  $K_{31}$ . Наконец, пучок, соединяющий  $F_{23}$  и  $K_{31}$ , пересекает  $E_1 \oplus E_2$  по искомому  $L_{12}$ . Затем для этих цепочек выписываются по (28) и (29) соотношения для изометрии (см. рисунок) [9, 12].

ТЕОРЕМА 5. В секторных пространствах формальной размерности три и выше клиффордова алгебра индуцированных гомотетий является телом, т.е. или полем  $\mathbb{R}$  вещественных чисел, или полем  $\mathbb{C}$  комплексных чисел, или телом  $\mathbb{H}$  кватернионов.

Этот эффект можно использовать для координации системы атомов секторного пространства аналогично тому, как это было сделано в теореме 3.

ТЕОРЕМА 6. Совокупность всех атомов сектора конечномерной  $\mathbb{R}$ -логики формальной размерности три и выше изоморфна связке всех прямых многомерного или вещественного, или комплексного, или кватернионного гильбертова пространства.

СЛЕДСТВИЕ. Йордановы логики в  $\mathbb{C}^n$  совпадают с  $\mathbb{R}$ -логиками.

10. Как было указано в п. 1, конечномерные некоммутативные распределения вероятностей задаются нормированными неотрицательными эрмитовыми операторами  $\Phi$ :

$$\Phi \geq 0, \quad \text{tr } \Phi = 1. \tag{30}$$

Их удобно погружать в линейное пространство зарядов — всех эрмитовых операторов со следовой нормой — суммой модулей всех своих собственных чисел. Как известно, каждый эрмитов оператор  $\Gamma$  однозначно раскладывается в разность своих положительной и отрицательной частей:

$$\Gamma = \Gamma^+ - \Gamma^-, \quad |\Gamma| = |\Gamma^+| + |\Gamma^-|. \tag{31}$$

ОПРЕДЕЛЕНИЕ 7. Положительная и отрицательная части эрмитова оператора  $\Gamma$ , а также его носитель  $G$  задаются через спектральное разложение оператора  $\Gamma$ :

$$\Gamma = \bigoplus_{\lambda_j} E_j, \quad G = \bigoplus_{\lambda_j \neq 0} E_j, \tag{32}$$

$$\Gamma^+ = \bigoplus_{\lambda > 0} \lambda_j E_j, \quad \Gamma^- = - \bigoplus_{\lambda < 0} \lambda_j E_j. \tag{33}$$

ОПРЕДЕЛЕНИЕ 8. Назовем *расщепом* любое замкнутое  $\mathbb{R}$ -линейное подпространство эрмитовых операторов, вместе с каждым  $\Gamma$  содержащее также  $\Gamma^+$ .

ТЕОРЕМА 7. Носители операторов расщепления образуют йорданову логику. Каждый атом этой логики несет единственный оператор распределения

вероятностей, который инвариантен относительно изометрий атома, индуцируемых секторной подлогикой по схеме теоремы 4.

Доказательство этой теоремы использует тот факт, что вместе с каждым двумя эрмитовыми операторами  $\Phi$  и  $\Gamma$  расщеп содержит  $\varepsilon^{-1}[\Phi - \varepsilon\Gamma]^-$ .

ЛЕММА 10. Если ядерные эрмитовы операторы  $\Phi$  и  $\Gamma$  неотрицательны и  $F$  — носитель оператора  $\Phi$ , то

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^{-1}[\Phi - \varepsilon\Gamma]^- = (1 - F)\Gamma(1 - F). \quad (34)$$

Эта лемма легко доказывается интегрированием по контуру в конечномерном случае [14]. Использование (34) позволяет расщепить шаг за шагом носители на атомы, которые могут в силу (34) нести только один неотрицательный нормированный заряд.

Теорема 7 позволяет дать описание всех вероятностных состояний, стационарных относительно марковского отображения — некоммутативного аналога конечной цепи Маркова, так как все стационарные операторы обязательно образуют расщепы (см. [14, 15], где рассмотрен также счетномерный случай и построен некоммутативный аналог структуры Деблина–Колмогорова).

Детали описания вытекают из принципа Веддерберна и теории спинорных представлений (см. [13, 16]). В соответствии с разложением (12) логики  $\Lambda$  расщеп  $\Sigma$  каждый оператор  $\Psi \in \Sigma$  разложим в ортосумму по секторам:  $\Psi = \oplus \Psi_j$ ,  $\Psi_j = Y_j \Psi Y_j$ . Ниже в описании  $\Psi$  латинскими буквами обозначены числа и операторы, жестко определяемые расщепом, греческими — произвольные. Если сектор  $Y_j$  одноатомный, то  $Y_j = \mathbb{C}^{t(j)}$ ,  $\Psi_j = \lambda Q_j$ ,  $Q_j$  — оператор вероятностей на  $\mathbb{C}^{t(j)}$ ,  $\lambda \in \mathbb{R}$ . Если подлогика  $Y_j \Lambda Y_j$  нейманова размерности  $k$ , то  $Y_j = \mathbb{C}^k \otimes \mathbb{C}^{t(j)}$  и, следовательно,  $\Psi_j = \Phi \otimes Q_j$ . В йордановом секторе  $Y_j$  над кватернионами  $\mathbb{H}$  или числами  $\mathbb{C}l_{2\nu}$  имеем при  $N = 2^\nu$  ( $\nu = 2$  для  $\mathbb{H}$ )  $Y_j = \mathbb{C}^k \otimes \mathbb{C}^N \otimes \mathbb{C}^{t(j)}$  и в соответствующем базисе матрица  $\Psi_j = \Phi \otimes Q_j$ , где  $\Phi$  — произвольная эрмитова матрица размерности  $k \times k$  над этими числами с каноническим (см. [13]) представлением мнимых единиц матрицами размерности  $N = 2^\nu$  (в клиффордовом случае обязательно  $k = 2$ ). В йордановых секторах над  $\mathbb{C}$  или над  $\mathbb{C}l_{2\nu+1}$  имеем  $T_j = \mathbb{C}^k \otimes \mathbb{C}^N \otimes \mathbb{C}^{t(j)} \otimes \mathbb{C}^k \otimes \mathbb{C}^N \otimes \mathbb{C}^{s(j)}$ ,  $\Psi_j = p_j \Phi \otimes Q_j \otimes p'_j \Phi' \otimes R_j$ ,  $p_j + p'_j = 1$ , матрица  $\Phi$  (и  $\Phi'$ ) — одна и та же эрмитова матрица над числами Клиффорда, интерпретируемая по первому (соответственно по второму) их каноническим представлениям, отличающимся знаками у интерпретации мнимых единиц (см. [13, 15]).

11. Ортопроекторы  $P(\varphi)$  на атомы  $\mathbb{R}$ -пучка в силу (8) и (14) линейно выражаются через любые различные три из них. Согласно теореме 7 эта связь сохраняется для минимальных состояний.

ЛЕММА 11. Если  $\varphi$  — угловая координата в пучке (14), то для любых четырех вероятностей  $\Psi\{P(\varphi)\} = \text{tr } \Psi P(\varphi)$  событий  $P(\varphi)$

$$\det \begin{vmatrix} 1 & \cos 2\varphi_1 & \sin 2\varphi_1 & \Psi\{P(\varphi_1)\} \\ \dots & \dots & \dots & \dots \\ 1 & \cos 2\varphi_4 & \sin 2\varphi_4 & \Psi\{P(\varphi_4)\} \end{vmatrix} = 0, \quad (35)$$

что обобщает соотношение Малюса в теории поляризованного света [17].

Так как в секторах клиффордова типа использовать трехмерную теорему Глисона [18] нельзя, то (35) не вытекает, вообще говоря, из условия аддитивности вероятности на логике и при задании вероятностей только на логике событий должно постулироваться дополнительно.

**ТЕОРЕМА 8.** *Всякая нормированная неотрицательная аддитивная  $\Psi\{F \oplus \oplus G\} = \Psi\{F\} + \Psi\{G\}$  оценка  $\Psi$  на конечномерной йордановой логике, удовлетворяющая также соотношению (35), задается эрмитовым оператором состояния на порождаемой логикой йордановой алгебре наблюдаемых.*

Доказательство использует минимаксные свойства собственных чисел эрмитовых матриц с элементами из  $\mathbb{R}, \mathbb{C}, \mathbb{H}$  или  $Cl$ . Таким образом, результаты п. 10 и 11 дополняют и уточняют рассуждения [19].

### СПИСОК ЛИТЕРАТУРЫ

1. Birkhoff G., Neumann J.von. The logic of quantum mechanics // Ann. Math. 1936. V. 37. P. 823–835.
2. Jordan P., Neumann J.von, Wigner E. On an algebraic generalization of the quantum mechanical formalism // Ann. Math. 1934. V. 35. P. 29–64.
3. Varadarajan V.S. Geometry of quantum theory. — N.Y.: Van Nostrand, 1968.
4. Нейман И. фон. Обобщение математического аппарата квантовой механики методами абстрактной алгебры. Ч. 1 // Мат. сб. 1936. Т. 1, № 4. С. 415–484.
5. Etmch G.G. Algebraic methods in statistical mechanics and quantum field theory. — N.Y.: Wiley, 1972.
6. Wong Young-Chow. Isoclinic  $n$ -planes in Euclidean  $2n$ -space, Clifford parallels in elliptic  $(2n - 1)$ -space, and the Hurwitz matrix equations // Mem. Amer. Math. Soc. 1961. V. 41. P. 1–112.
7. Широков П.А. Тензорное исчисление. — 2-е изд. — Казань: Изд-во КГУ, 1962.
8. Jordan C. Essai sur la geometric a “ $n$ ” dimensions // Bull. Soc. math. France. 1874-75. V. 3. P. 103–174; Oeuvres. P.: Gauthier-Villars et Blanchard, 1962. V. 3. P. 79–149.
9. Морозова Е.А., Ченцов Н.Н. Элементарные йордановы логики: Препр. ИПМ АН СССР № 113. — М., 1975. 40 с.
10. Делоне Б.Н., Райков Д.А. Аналитическая геометрия. Т. 1. — М.-Л.: Гостехиздат, 1948.
11. Морозова Е.А., Ченцов Н.Н. Некоммутативные квантовые логики. — М., 1981. — 28 с. — (Препр. / ИПМ АН СССР; № 57).
12. Морозова Е.А., Ченцов Н.Н. К теореме Йордана–фон Неймана–Вигнера. — М., 1975. — 70 с. — (Препр. / ИПМ АН СССР; № 129).
13. Chevalley C. Theory of Lie groups. Vol. 1. — Princeton: Univ. press, 1946.
14. Морозова Е.А., Ченцов Н.Н. Алгебра ограниченных гармонических функций счетной цепи Маркова (некоммутативная теория). — М., 1981. — 28 с. — (Препр. / ИПМ АН СССР; № 1).
15. Морозова Е.А., Ченцов Н.Н. Структура семейства стационарных состояний квантовой цепи Маркова. — М., 1976. — 55 с. — (Препр. / ИПМ АН СССР; № 130).
16. Желобенко Д.П. Компактные группы Ли и их представления. — М.: Наука, 1970.
17. Shurkliff W.A. Polarized light. — Cambridge: Harvard Univ. press, 1962.
18. Gleason A.M. Measures on the closed subspaces of a Hilbert space // J. Rat. Mech. Anal. 1957. V. 6. P. 885–894.
19. Gunson J. Physical states on quantum logics. I // Ann. Inst. H. Poincare. 1972. V. 17, № 4. P. 295–311.

РАЗДЕЛ V

**ЗАДАЧИ МЕХАНИКИ  
И ФИЗИКИ**

---

**ОПТИМАЛЬНАЯ СТРАТЕГИЯ  
ПРИ КОРРЕКТИРОВАНИИ<sup>\*)</sup>**

Настоящая заметка посвящена вопросу построения оптимальной стратегии корректирования аппарата, движущегося вблизи номинальной траектории. В отличие от работ [1–3], где рассматриваются стратегии с моментами коррекции, не зависящими от результатов траекторных измерений, здесь рассматриваются гибкие стратегии, в которых моменты коррекции и величины корректирующих импульсов назначаются во время полета по результатам проведенных траекторных измерений. Вместо критерия минимума математического ожидания расхода топлива взят прямой критерий максимума вероятности попадания в заданную область пространства корректируемых параметров при данном ограниченном запасе топлива.

1. В основе рассмотрения лежит следующая линеаризованная модель. Движение аппарата в окрестности номинальной траектории описывается системой дифференциальных уравнений

$$\dot{\mathbf{h}}(t) = R(t)\mathbf{h}(t) + R_1(t)\mathbf{u}(t),$$

где  $\mathbf{h}(t)$  — шестимерный вектор фазовых отклонений фактической траектории от номинальной,  $\mathbf{u}(t)$  — управление, состоящее из конечного числа управляющих импульсов. Управляющий импульс меняет скорость аппарата, оставляя неизменными его координаты. Предполагается, что ошибки исполнения отсутствуют. Одномерное пространство величин  $\delta$ , являющихся линейными комбинациями компонент вектора  $\mathbf{h}(T)$ , называется *пространством корректируемых параметров* или *пространством промахов*. В этом пространстве выбран интервал допустимых промахов  $A = (-\delta_0, \delta_0)$ . Считается, что коррекция прошла успешно, если в момент  $T$  происходит событие

$$\{\delta \in A\}. \tag{1}$$

В фиксированные моменты времени  $t_1, \dots, t_N$  проводятся траекторные измерения. Измеряются отклонения  $\alpha_i$  некоторых величин от номинальных, эти

---

<sup>\*)</sup> Докл. АН СССР. 1967. Т. 175. С. 1.47-50 (совм. с Д.Е. Огоцимским, В.А. Рясным).

отклонения линейно зависят от ошибок выведения  $\mathbf{h}(0)$  и управления на интервале времени, предшествующем измерению. Измерения проводятся с ошибками  $\Delta_i$ . Результат измерения записывается в виде

$$x_i = b_i(\mathbf{u}[0, t_i]) + \mathbf{I}_i \cdot \mathbf{h}(0) + \Delta_i = \alpha_i + \Delta_i,$$

где  $b_i$  — член, характеризующий влияние на  $\alpha_i$  управления на интервале  $[0, t_i]$ ;  $\mathbf{I}_i$  — вектор, связанный с моментом  $t_i$ . Пространство  $X$  совокупностей  $x = (x_1, \dots, x_N)$  называется *пространством выборок*. Предполагается, что на пространстве ошибок выведения и ошибок измерений (пространство  $\Omega$  элементарных исходов) задана гауссова вероятность.

*Стратегия* — правило, позволяющее для каждой выборки  $x$  наблюдений назначать моменты коррекций и корректирующие импульсы. Решение о каждой коррекции принимается на основании фактических траекторных измерений, числа и точности будущих траекторных измерений, имеющегося запаса топлива и числа коррекций, проведение которых предполагается в будущем. В заметке рассматриваются стратегии с одной и двумя коррекциями. Стратегия с одной коррекцией определяется разбиением пространства выборок на  $N$  цилиндрических множеств  $B_1, \dots, B_N$  и заданием на каждом  $B_i$  функции времени  $t_i \leq \tau(\mathbf{x}_i) < t_{i+1}$  и управляющей функции  $|\mathbf{q}(\mathbf{x}_i)| \leq W_0$ , где  $W_0$  — запас топлива на коррекцию,  $\mathbf{x}_i = (x_1, \dots, x_i)$ . Стратегия в двухразовой коррекции определяется разбиением пространства выборок на цилиндрические множества  $B_{ij}$ ,  $i = 1, 2, \dots, N-1$ ,  $i < j \leq N$ , на которых заданы функции времени и управляющие функции  $\mathbf{q}_1(\mathbf{x}_i)$  для первой коррекции  $t_i \leq \tau_1(\mathbf{x}_i) < t_{i+1}$  и  $\mathbf{q}_2(\mathbf{x}_j, i)$  для второй коррекции  $t_j \leq \tau_2(\mathbf{x}_j, i) < t_{i+1}$  так, что  $|\mathbf{q}_1| + |\mathbf{q}_2| \leq W_0$ .

Каждой стратегии  $\nu$  соответствует определенная вероятность  $P(\nu)$  события (1). Стратегия  $\nu^*$  оптимальна в данном классе, если для любой другой стратегии  $\nu$  из этого класса  $P(\nu) \leq P(\nu^*)$ . Задача состоит в построении оптимальных стратегий в классах  $E_1$  и  $E_2$  стратегий с одной и двумя коррекциями.

**2.** Выделим из классов  $E_1$  и  $E_2$  так называемые полные подклассы  $\Pi_1$  и  $\Pi_2$  стратегий, обладающих преимуществом в способе задания и содержащих оптимальную стратегию. Рассмотрим функцию влияния  $\theta(t)$ , показывающую, какой максимальный сдвиг в пространстве промахов можно получить коррекцией в момент  $t$ , располагая единичным запасом топлива. Направление импульса, приводящего к максимальному сдвигу, обозначим  $\vec{\theta}(t)$ . Пусть  $\theta(t)$  непрерывна. Через  $\theta_i$  обозначим максимум  $\theta(t)$  в  $[t_i, t_{i+1})$ . Выберем моменты времени  $\tau_i$ , удовлетворяющие следующим условиям:  $t_i \leq \tau_i < t_{i+1}$ ,  $\theta(\tau_i) = \theta_i$ ,  $\theta(t) < \theta_i$  при  $t > \tau_i$ . Таких моментов не больше, чем один на каждом интервале  $[t_i, t_{i+1})$ , и если функция влияния не монотонна, то на некоторых интервалах они могут отсутствовать.

Классу  $\Pi_m$ ,  $m = 1, 2$ , по определению принадлежат те и только те стратегии из  $E_m$ , которые разрешают коррекцию лишь в моменты  $\tau_i$  в направлении  $\vec{\theta}(\tau_i)$ .

**ЛЕММА.** Для всякой стратегии  $\nu \in E_m$  существует стратегия  $\rho \in \Pi_m$  такая, что  $P(\nu) = P(\rho)$ , и требующая меньшего расхода топлива.

На основании леммы  $\Pi_m$  образует полный класс, которым мы и ограничимся при дальнейших рассуждениях.

**3.** Разбиение пространства выборок вместе с совокупностью функций момен-

та и величины импульса представляет собой способ задания стратегии как функции на пространстве элементарных исходов. Рассматриваемые классы стратегий на  $\Omega$  удобно задать с помощью другого пространства выборок  $Z = \{z_1 \dots z_N\} = \{z\}$ , координаты которого являются разностями прогнозов промаха при последовательных измерениях в условиях отсутствующего управления. Под *прогнозом промаха*  $z(i)$  в момент  $t_i$  в условиях отсутствующего управления понимается математическое ожидание промаха при известных  $x_1, \dots, x_i$ , если на  $[0, t_i]$  управление равно нулю. Если же на  $[0, t_i]$  управление не равно нулю, то в определении прогноза  $x_1, \dots, x_i$  следует заменить на  $y_1, \dots, y_i$ , где  $y_i = \mathbf{I}_i \cdot \mathbf{h}(0) + \Delta_j = x_j - b_j(\mathbf{u}[0, t_j])$ . Доказано, что рассматриваемый класс стратегий может быть задан с помощью  $Z$ .

4. Рассмотрим одноразовую коррекцию. Чтобы упростить обозначения, будем считать, что возможные для проведения коррекции моменты  $\tau_i$  имеются на каждом интервале  $[t_i, t_{i+1})$ . Обобщение на случай меньшего числа возможных моментов коррекции очевидно.

Для стратегии с одной коррекцией выпишем выражение для вероятности успеха коррекции:

$$P(\nu) = \sum_k \int_{B_k} \left[ \int_A \frac{1}{\sqrt{2\pi} \sigma_k} \exp -\frac{1}{2\sigma_k^2} (\delta - z(k) - q(\mathbf{z}_k)\theta_k)^2 d\delta \right] d\mathcal{P}(\mathbf{z}_k),$$

где  $\sigma_k^2 = \sigma_0^2 - Mz_1^2 - \dots - Mz_k^2$  и  $\sigma_0^2$  — априорная дисперсия промаха. В квадратных скобках стоит вероятность успеха  $\varphi(z(k), q(\mathbf{z}_k))$  при известных  $\mathbf{z}_k, q(\mathbf{z}_k)$ .

В качестве управляющей функции в момент  $\tau_i$  рассмотрим

$$Q(\mathbf{z}_k) = \begin{cases} -\frac{z(k)}{\theta_k} & \text{при } |z(k)| \leq W_0\theta_k, \\ -\text{sign } z(k) \cdot W_0 & \text{при } |z(k)| > W_0\theta_k. \end{cases} \quad (2)$$

Эта функция максимизирует значение условной вероятности успеха коррекции при каждой выборке  $\mathbf{z}_k$ . Значит, при любом разбиении  $\{B_k\}$  пространства  $Z$  в качестве управляющих функций всегда следует брать  $Q(\mathbf{z}_k)$ . Оптимальное разбиение  $\{B_{k \text{ опт}}\}$  можно найти методом динамического программирования. Обозначим  $\psi(\mathbf{z}_k) = \varphi(z(k), Q(\mathbf{z}_k))$ ,  $r(\mathbf{z}_N) = \psi(\mathbf{z}_N)$ ,  $r(\mathbf{z}_k) = \max\{\psi(\mathbf{z}_k), \int r(\mathbf{z}_{k+1}) d\mathcal{P}(z_{k+1})\}$  и введем множества  $C_k = \{z: \psi(\mathbf{z}_k) > \int r(\mathbf{z}_{k+1}) d\mathcal{P}(z_{k+1})\}$ . Тогда  $\{B_{k \text{ опт}}\}$  записывается в виде  $C_1, \overline{C_1} \cap C_2, \dots, (\overline{C_1} \cup \dots \cup \overline{C_{N-1}})$ . Оптимальная управляющая функция (2) в момент  $\tau_k$  при фиксированном запасе топлива  $W_0$  зависит только от прогноза, а не от всей системы измерений  $z_1, \dots, z_k$ . Значит,  $\psi(\mathbf{z}_k)$  зависит от прогноза. Формулы для  $r(\mathbf{z}_N)$  и  $r(\mathbf{z}_k)$  позволяют заключить по индукции, что и  $r(\mathbf{z}_k)$  зависит только от

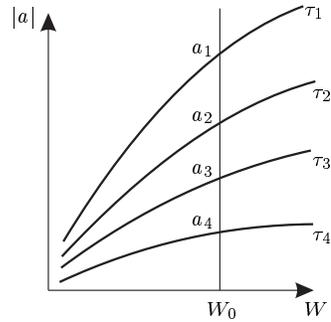


Рис. 1. Одноразовая коррекция. Область  $D_k$  для каждого момента  $\tau_k$  есть часть прямой  $W = W_0$  при  $|a| > a_k$

прогноза. Обозначим  $\psi(z_k)$  и  $r(z_k)$  через  $\lambda_k(a)$  и  $\mu_k(a)$ , где  $a = z(k)$ , введем на числовой оси множества

$$D_k = \left\{ a: \lambda_k(a) > \int \mu_{k+1}(a + z_{k+1}) dP(z_{k+1}) \right\}.$$

Тогда, как следует из выражения для  $C_k$ , множеству  $C_k$  принадлежат те выборки из  $Z$ , сумма первых  $k$  элементов которых есть число из  $D_k: C_k = \{z: z(k) \in D_k\}$ . Процесс принятия решения состоит в следующем: 1) коррекция назначается в момент  $\tau_k$ , когда прогноз первый раз попадает в область  $D_k$ ; 2) величина корректирующего импульса назначается согласно формуле (2). Иллюстративный график приведен на рис. 1.

Рассмотрим случай почти точных измерений, когда  $\sigma_0^2 - Mz_1^2 \ll \sigma_0^2$ . Это условие означает, что случайное смещение прогноза в момент  $\tau_i, i > 1$  от прогноза в момент  $\tau_1$  мало в масштабе  $\sigma_0$ , хотя оно может быть большим в масштабе интервала  $A$  допустимых промахов. Функция  $\lambda_k(a)$  в условиях почти точных измерений имеет вид почти ступеньки. Область размытого скачка находится в окрестности точки  $a = W_0\theta_k$  и имеет величину порядка  $3\sigma_k$ , малую в масштабе  $3\sigma_0$ . Легко определяются области  $D_k$  и функции  $\mu_k(a): D_k \approx \{a: a > W_0\theta_{k+1}\}$  с точностью  $3\sigma_k$ ,

$$\mu_k(a) = \begin{cases} \Phi(\delta_0/\sigma_k) - \Phi(-\delta_0/\sigma_k) & \text{при } W_0\theta_{i+1} \leq |a| < W_0\theta_i, \quad i = k+1, \dots, N, \\ \lambda_k(a) & \text{при } |a| > W_0\theta_{k+1}, \end{cases}$$

где  $\theta_{N+1} = 0, \Phi(x)$  — функция распределения стандартного нормального закона. Оптимальная вероятность успеха вычисляется по формуле  $P(v^*) = \int \mu_1(a) dP(a)$  и приближенно равна

$$\sum_1^N \left[ \Phi\left(\frac{\delta_0}{\sigma_i}\right) - \Phi\left(-\frac{\delta_0}{\sigma_i}\right) \right] \left[ \Phi\left(\frac{W_0\theta_i}{\sqrt{Mz_1^2}}\right) - \Phi\left(-\frac{W_0\theta_{i+1}}{\sqrt{Mz_1^2}}\right) \right].$$

**5.** Рассмотрим двухразовую коррекцию. Вопрос о построении оптимальной стратегии сводится к вопросу о построении оптимальной стратегии в подклассе  $U_k$  стратегий с двумя коррекциями, для которых момент первой коррекции с вероятностью 1 равен  $\tau_k$ .

**ТЕОРЕМА.** *В подклассе  $U_k$  стратегий с двумя коррекциями и с фиксированным моментом первой коррекции существует оптимальная.*

Доказательство теоремы связано с приближением в пространстве стратегий.

Можно показать также, что среди оптимальных есть стратегия, в которой управляющие функции для первого импульса  $Q_1(a)$  зависят лишь от прогноза промаха  $a = z(k)$ . Для такой стратегии на прямой вводятся множества, зависящие от запаса топлива,  $D'_l(W)$  и  $D'_k(W)$ , которые, как и в случае однократной коррекции определяют оптимальное разбиение  $\{B_{k \text{ опт}}\}$  и  $\{B_{kl \text{ опт}}\}$  пространства  $Z$ . Способ поведения при двухразовой коррекции состоит в следующем. Суммируются координаты  $z$ . Как только происходит событие  $\{z(k) \in D'_k(W_0)\}$ , проводится первая коррекция импульсом величины  $Q_1(z(k))$ . Производятся последующие измерения и назначается вторая коррекция, как только происходит событие  $\{a' \in D'_l(W_0 - |Q_1(z(k))|)\}$ , где  $a' = z(k) + Q_1(z(k))\theta_k + z_{k+1} + \dots + z_l$ .

Функция управления для второго импульса получается, если в (2) значение  $W_0$  заменить на  $W_0 - |Q_1(z(k))|$ . Расчеты показывают, что если измерения не являются почти точными, то оптимальная управляющая функция для первого импульса в момент  $\tau_k$  по абсолютной величине меньше функции (2), т.е. имеет место недокоррекция. При почти точных измерениях недокоррекция исчезает. В этом случае можно предложить много способов, близких к оптимальному. Простейший состоит в том, что после первого измерения сообщается импульс  $Q_1 = z(1)/\theta_1$ . Вторая коррекция проводится в момент  $\tau_N$ :  $Q_2 = z(N)/\theta_N$ . Вероятность успеха при такой стратегии приближенно равна

$$\left[ \Phi\left(\frac{\delta_0}{\sigma_N}\right) - \Phi\left(-\frac{\delta_0}{\sigma_N}\right) \right] \left[ \Phi\left(\frac{W_0\theta_1}{\sqrt{Mz_1^2}}\right) - \Phi\left(-\frac{W_0\theta_1}{\sqrt{Mz_1^2}}\right) \right].$$

Ошибка в вероятности успеха по сравнению с оптимальной стремится к нулю с ростом точности измерений.

#### СПИСОК ЛИТЕРАТУРЫ

1. *Breakwell J.V., Tung F., Smith R.R.* // AIAA J. 1965. V. 3, № 5. P. 807.
2. *Tung F.* // IEEE. Trans. on Automatic Control, AC-10. 1965. № 3. P. 328.
3. *Ярошевский В.А., Парышева Т.В.* // Космические исследования. 1965. Т. 3, № 6. С. 826.
4. *Рясин В.А.* // Теория вероятностей и ее применение. 1966. Т. 11, № 4. С. 708.

## ВЫЧИСЛЕНИЕ УГЛОВЫХ КОЭФФИЦИЕНТОВ ИЗЛУЧЕНИЯ В ОСЕСИММЕТРИЧНОЙ ГЕОМЕТРИИ<sup>\*</sup>)

При проектировании печей, высокотемпературного химического оборудования, высокотемпературных устройств для преобразования энергии и криогенных систем необходимо рассчитывать теплообмен излучением.

Из-за сложности радиационного теплообмена и отсутствия точных значений излучательных характеристик поверхностей обычно при расчете рассматриваются модели и оболочки с простыми свойствами поверхностей.

При расчете теплообмена излучением между диффузно излучающими и диффузно отражающими поверхностями, разделенными диатермической средой, необходимо вычислить угловые коэффициенты излучения, которые определяют долю энергии, передающуюся с одной поверхности на другую.

Опубликовано значительное количество работ, в которых угловые коэффициенты вычисляются аналитическим способом для различных конфигураций

<sup>\*</sup>) Инженерно-физический журнал. 1978. Т. 34, № 2. С. 306–312. (совм. с Г.В. Думкиной, Н.С. Шойдиной).

(см. [1–4] и др.). В данной работе предлагается алгоритм расчета угловых коэффициентов излучения на ЭВМ.

В случае когда плотность потока излучения равномерно распределена по поверхности, угловой коэффициент между поверхностями  $p$  и  $q$  —

$$\varphi_{pq} = \frac{\int_{A_p} \int_{A_q} \frac{\cos \beta_i \cos \beta_j}{\pi L_{ij}^2} dA_i dA_j}{\int_{A_p} dA_i}, \quad (1)$$

где  $A_p, A_q$  — площади поверхностей  $p, q$ ;  $L_{ij}$  — расстояние между элементами поверхностей  $i, j$ ;  $\beta_i, \beta_j$  — углы между внутренними нормальными к элементам поверхности и линией, соединяющей эти элементы.

Область интегрирования распространяется по тем частям поверхностей  $p, q$ , которые непосредственно видят друг друга.

Вычислим угловые коэффициенты в осесимметричной системе, состоящей из полости и внутренних тел (рис. 1). В качестве областей  $p, q$  возьмем пояски поверхностей  $r = r(z)$ , вырезаемые плоскостями  $z = \text{const}$ . Каждая область  $p(q)$

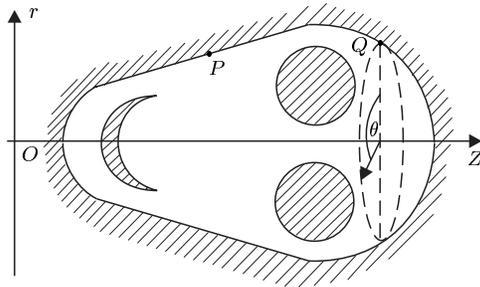


Рис. 1. Осесимметричная система:  $z$  — координата вдоль оси вращения;  $r$  — расстояние до оси;  $\theta$  — угол поворота вокруг оси

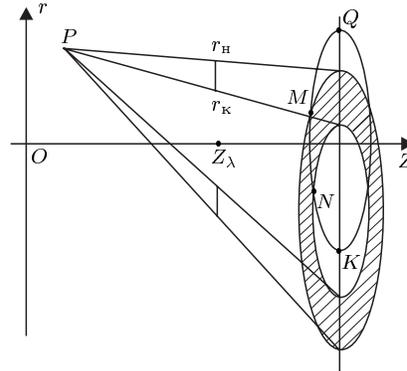


Рис. 2. Область видимости точек  $P$  и  $Q$ :  $\overset{\frown}{MQ}, \overset{\frown}{NK}$

разбивается на более мелкие пояски  $i(j)$ . Для каждой пары  $i, j$  вычисляются искомые интегралы, затем для каждой пары  $p, q$  эти интегралы суммируются по всем  $i$  номерам мелких поясков, на которые разбита область  $p$ , и по всем  $j$  номерам, на которые разбита область  $q$ .

Для вычисления интеграла применяются следующие приближенные формулы:

$$\frac{1}{A_i} \int_{A_i} \int_{A_j} f_{ij} dA_i dA_j = \frac{A_j}{2\pi^2} \int_0^{\theta_{ij}} f_{ij}(r_i, z_i, 0, r_j, z_j, \theta) d\theta. \quad (2)$$

Справедлива теорема: область видимости точек  $P(r_i, 0, z_i)$  и  $Q(r_j, \theta, z_j)$ , определяемая произвольным телом вращения, в осесимметричной конфигурации состоит не более чем из двух дуг в  $[0, \pi]$ .

Рассмотрим мешающее тело, образованное вращением отрезка  $r_n r_k$  вокруг оси  $z$  (рис. 2). Возможны следующие случаи:

- а)  $r_n, r_k$  — любые конечные значения  $r \neq 0$ ;

б)  $r_k = 0$ ,  $r_n > r_k$ ,  $r_n = \text{const}$ ;

в)  $r_n \rightarrow \infty$ ,  $r_k \neq 0$ .

Построим два конуса с вершиной в точке  $P$  и направляющими, проходящими через точки  $r = r_n$ ,  $z = z_\lambda$  (I конус),  $r = r_k$ ,  $z = z_\lambda$  (II конус) (случай а)). При повороте точки  $Q$  вокруг оси видимость точек  $P$  и  $Q$  имеется, когда луч  $PQ$  находится вне конуса I или внутри конуса II. Отсюда область видимости точек  $P$  и  $Q$  для  $\theta \in [0, \pi]$  состоит не более чем из двух дуг:  $[0, \theta_{ij}^k]$ ,  $[\theta_{ij}^n, \pi]$ .

При  $r_k = 0$  (случай б)) область видимости состоит из одной дуги  $[0, \theta_{ij}^k]$ , при  $r_n \rightarrow \infty$  (случай в)) область видимости равна  $[\theta_{ij}^n, \pi]$ .

Мешающее тело, лежащее на оси (рис. 3, а), можно представить как совокупность большого количества достаточно тонких цилиндров (случай б)),

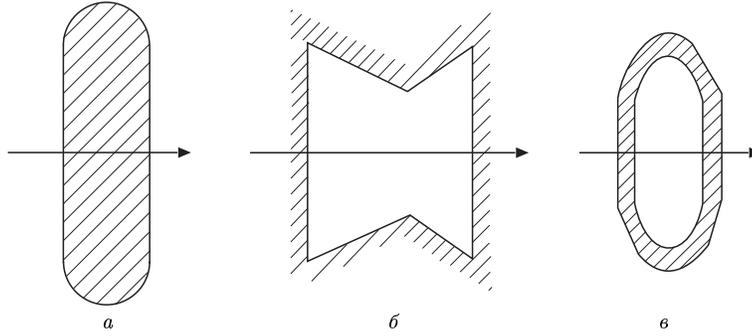


Рис. 3. Различные типы мешающих тел

для которых  $r_{k\lambda} = 0$ ,  $r_{n\lambda} = \text{const}(\lambda)$ , где  $\lambda$  — номер цилиндра. Область видимости  $\Delta_{ij}$ , определяемая данным мешающим телом,  $\Delta_{ij} = [0, \theta_{ij}^k]$ , где  $\theta_{ij}^k = \min_\lambda \theta_{ij}^k \lambda$ . Аналогично область видимости, определяемая поверхностью осесимметричной полости (рис. 3, б), равна  $[\theta_{ij}^n, \pi]$ , где

$$\theta_{ij} = \max_\lambda \theta_{ij}^n \lambda.$$

Область видимости, определяемая мешающим телом типа тора (рис. 3, в), равна в общем случае

$$\Delta_{ij} = [0, \theta_{ij}^k] \cap [\theta_{ij}^n, \pi],$$

где

$$\theta_{ij}^k = \min_\lambda \theta_{ij}^k \lambda, \quad \theta_{ij}^n = \max_\lambda \theta_{ij}^n \lambda.$$

Рассмотрим системы, для которых мешающими телами являются тела, лежащие на оси и поверхности осесимметричной полости. Область видимости для любой пары точек  $P$  и  $Q$  состоит из одной дуги  $[\theta_{ij}^n, \theta_{ij}^k]$ .

Для того чтобы точки  $P(r_i, 0, z_i)$  и  $Q(r_j, \theta, z_j)$  видели друг друга, необходимо и достаточно, чтобы: 1) углы  $\beta_i, \beta_j$  были острыми; 2) отрезок  $PQ$  не встречал никаких мешающих тел.

Положим, что поверхности мешающих тел являются либо цилиндрами, либо конусами, либо сферами. Эти поверхности можно записать общим уравнением

$$r_\lambda^2 = k_\lambda(z - \xi_\lambda) + R_\lambda^2, \quad (3)$$

где  $\lambda$  — номер поверхности;  $\xi_\lambda$  — координата центра шара или вершина конуса;  $k_\lambda = -1$  для сферы;  $k_\lambda = \operatorname{tg}^2 \psi$  для конуса;  $\psi + \pi/2$  — угол между внутренней нормалью к поверхности и положительным направлением оси вращения;  $k_\lambda = 0$  для цилиндра;  $R_\lambda = 0$  для конуса;  $R_\lambda$  — радиус сферы, цилиндра.

Прямая  $PQ$  и мешающая поверхность (3) определяют координату  $z$  точки пересечения

$$az^2 - 2bz + c = 0, \quad (4)$$

где

$$\begin{aligned} a &= r_i^2 + r_j^2 - 2r_i r_j \cos \theta - k_\lambda (z_j - z_i)^2 = m + h \cos \theta, \\ b &= r_i^2 z_j + r_j^2 z_i - r_i r_j (z_i + z_j) \cos \theta - k_\lambda \xi_\lambda (z_j - z_i)^2 = l + t \cos \theta, \\ c &= r_i^2 z_j^2 + r_j^2 z_i^2 - 2r_i r_j z_i z_j \cos \theta - (k_\lambda \xi_\lambda^2 + R_\lambda^2) (z_j - z_i)^2 = s + g \cos \theta. \end{aligned}$$

Луч  $PQ$  пересекает мешающее тело, если корни уравнения (4) действительные и точка пересечения  $(r_p, z_p)$  лежит:

1) между  $P$  и  $Q$ , т.е.

$$z_i \leq z_p \leq z_j; \quad (5)$$

2) на поверхности мешающего тела, а не на его продолжении (мешающее тело ограничено по оси  $z$  координатами  $z_{н\lambda}, z_{к\lambda}$ ), т.е.

$$z_{н\lambda} \leq z_p \leq z_{к\lambda}. \quad (6)$$

Для определения области видимости выделяем среди мешающих тел те, которые влияют на величину угла видимости точек  $P$  и  $Q$ :

$$z_i \leq (z_{н\lambda} \vee z_{к\lambda}) \leq z_j. \quad (7)$$

Затем определяем интервал видимости по каждой мешающей поверхности независимо.

Схема определения области видимости точек  $P$  и  $Q$  ( $z_i \neq z_j$ ) по одной мешающей поверхности следующая:

- 1) определяем видимость между точками  $P$  и  $Q$  при  $\theta = 0$ ;
- 2) проверяем, видит ли точка  $P$  точку  $Q$  при  $\theta = \pi$ .

В результате проверки 1), 2) для мешающей поверхности возможен один из четырех случаев:

	I	II	III	IV
Видимость при $\theta = 0$	Есть	Нет	Есть	Нет
Видимость при $\theta = \pi$	Есть	Нет	Нет	Есть

В случае I область видимости точек  $P$  и  $Q$  —  $[0, \pi]$ , в II — равна нулю, в случае III —  $[0, \theta_\lambda^к]$ , в IV —  $[\theta_\lambda^н, \pi]$  ( $\theta_\lambda^к$  — угол, при котором прямая  $PQ$  при вращении точки  $Q$  вокруг оси  $z$  коснется мешающей поверхности).

Из уравнения касательной к мешающей поверхности находим

$$\cos \theta_\lambda^k = \begin{cases} \frac{R_\lambda^2 - \xi_\lambda^2 - z_i z_j + \xi_\lambda(z_i + z_j) - \sqrt{(r_i^2 - \rho_i^2)(r_j^2 - \rho_j^2)}}{r_i r_j} & \text{для сферы,} \\ \frac{\rho_i \rho_j - \sqrt{(r_i^2 - \rho_i^2)(r_j^2 - \rho_j^2)}}{r_i r_j} & \text{для конуса, цилиндра,} \end{cases} \quad (8)$$

$$\rho_v^2 = k_\lambda(z_v - \xi_\lambda)^2 + R_\lambda^2, \quad v = i, j.$$

Касание определяет угол видимости, если для координаты точки касания выполнены условия (5), (6).

Если касания нет, то определяем угол, при котором прямая  $PQ$  пройдет через край мешающей поверхности ( $z_{н\lambda}$  или  $z_{к\lambda}$ ):

$$\cos \theta_\lambda^k = \frac{2lz_{н\lambda} - s - mz_{н\lambda}^2}{hz_{н\lambda}^2 - 2tz_{н\lambda} + g}. \quad (9)$$

Если при повороте точки  $Q$  прямая  $PQ$  пройдет через оба края, то из двух полученных значений  $\theta_\lambda^k$  выбираем наименьшее;  $\theta_\lambda^н$  определяются из условия прохождения прямой  $PQ$  через край мешающей поверхности.

Рассматривая совокупность мешающих поверхностей (3), получаем последовательности:

$$\text{I: } \theta_1^н, \theta_2^н, \dots, \theta_n^н; \quad \text{II: } \theta_1^к, \theta_2^к, \dots, \theta_n^к;$$

где  $1, 2, \dots, n$  — номера мешающих поверхностей.

Область видимости для точек  $P$  и  $Q$  равна  $(\theta_{ij}^н, \theta_{ij}^к)$ , где

$$\theta_{ij}^н = \max_\lambda \theta_\lambda^н, \quad \theta_{ij}^к = \min_\lambda \theta_\lambda^к.$$

Если  $\theta_{ij}^н > \theta_{ij}^к$ , то точка  $P$  не видит точку  $Q$ .

Для случая  $z_i = z_j$  мешать могут только те тела, проекции которых на ось  $z$  содержат проекции точек  $P$  и  $Q$ .

Если точки  $P$  и  $Q$  не видят друг друга при  $\theta = 0$ , то область видимости равна 0. Если видимость есть при  $\theta = 0$  и  $\theta = \pi$ , то область видимости —  $[0, \pi]$ . Если видимость при  $\theta = \pi$  равна нулю, то угол видимости определяется из условия касания прямой  $PQ$  мешающей поверхности:

$$\cos \theta_\lambda^к = \frac{\rho_\lambda^2 - \sqrt{(r_i^2 - \rho_\lambda^2)(r_j^2 - \rho_\lambda^2)}}{r_i r_j}. \quad (9')$$

Подынтегральная функция в (2) есть

$$f_{ij} = \frac{1}{\pi} \frac{(A_{ij} + B_{ij} \cos \theta)(C_{ij} + D_{ij} \cos \theta)}{(E_{ij} + F_{ij} \cos \theta)^2},$$

коэффициенты  $A_{ij}$ ,  $B_{ij}$ ,  $C_{ij}$ ,  $D_{ij}$ ,  $E_{ij}$ ,  $F_{ij}$  вычисляются по значениям  $r_i$ ,  $z_i$ ,  $\psi_i$ ,  $r_j$ ,  $z_j$ ,  $\psi_j$ .

Интеграл  $I = \int_{\theta_{ij}^н}^{\theta_{ij}^к} f_{ij} d\theta$  вычисляется по формуле трапеции.

Для малых углов ( $\theta \leq 30^\circ$ ) и для  $z_i = z_j$  получено аналитическое выражение для интеграла.

Таблица угловых коэффициентов  $\varphi_{ij}$

Номер области	1	2	3	4	5	6	7
3	0	0	0	0	0	0,06688	0,27174
30	0	0	0	0,00061	0,0009	0	0
34	0	0	0	0	0	0	0
Номер области	8	9	10	11	12	13	14
3	0,1512	0,10993	0,08134	0,06119	0,00044	0,01463	0,01596
30	0,00051	0,00086	0,00155	0,00306	0,04347	0,04628	0,10031
34	0	0,02097	0,10796	0,16726	0,06494	0,11469	0,09496
Номер области	15	16	17	18	19	20	21
3	0,01578	0,01369	0,01121	0,00885	0,00712	0,00583	0,01827
30	0,19174	0,24544	0,18213	0,09302	0,04284	0,02047	0,02053
34	0,07791	0,06153	0,03957	0,01982	0,00961	0,00498	0,00417
Номер области	22	23	24	25	26	27	28
3	0,00849	0	0	0	0	0	0
30	0,0066	0	0	0	0	0	0
34	0	0	0	0	0	0	0
Номер области	29	30	31	32	33	34	35
3	0	0	0	0,02292	0	0	0,11609
30	0	0	0	0	0	0	0
34	0	0	0	0,14709	0,06548	0	0

В качестве иллюстрации в таблице приведены полученные значения угловых коэффициентов для системы, изображенной на рис. 4.

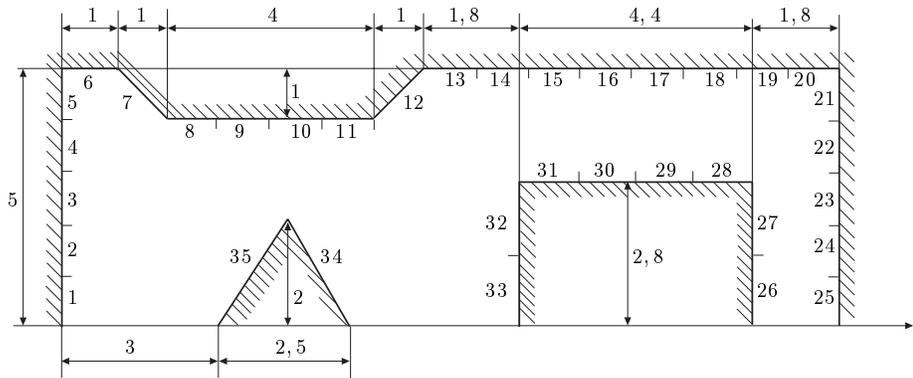


Рис. 4. Конфигурация рассчитываемой системы,  $z$  — ось вращения. Приведено разбиение на области. Размеры даны в сантиметрах

В системе 166 счетных точек. В областях 1–6, 8–11, 13–25, 28–31 по 2 счетные точки, в областях 7, 12, 26, 27, 32, 33 по 3, в областях 34, 35 по 5 счетных точек, разбиение равномерное по длине.

Как видно из таблицы, точность счета угловых коэффициентов, определяемая с помощью алгебры угловых коэффициентов, достаточно высокая.

$\sum_{j=1}^{35} \varphi_{3,j} = 1,0016$ ,  $\sum_{j=1}^{35} \varphi_{30,j} = 1,00032$ ,  $\sum_{j=1}^{35} \varphi_{34,j} = 1,000999$ . В углах (области 5–6, 20–21) точность счета несколько ниже.

Время счета задачи на машине БЭСМ-6 — 20 с.

При определении области видимости традиционным способом (область  $[0, \pi]$  разбивается на 50 интервалов равной длины и определяется видимость для каждого интервала) время счета возрастает  $\sim$  в 10–20 раз при той же точности расчета.

#### СПИСОК ЛИТЕРАТУРЫ

1. *Спэрроу Э.М., Сесс Р.Д.* Теплообмен излучением. — Л.: Энергия, 1971.
2. *Суринов Ю.А., Буглаев Е.К., Прохач Э.Е.* // Труды Краснодарского политех. ин-та. 1973. Вып. 43. С. 109.
3. *Сосновая Э.Г., Суринов Ю.А., Сосновыи Н.В.* // Труды Краснодарского политех. ин-та. 1973. Вып. 54. С. 85.
4. *Соулз, О'Брайен.* Теплопередача, № 3. — М.: Мир, 1972.

## МНОГОМОДОВАЯ ВИНТОВАЯ МГД-НЕУСТОЙЧИВОСТЬ<sup>\*)</sup>

### Введение

Настоящая работа посвящена исследованию эволюции винтовых возмущений в цилиндрических плазменных конфигурациях с винтовыми магнитными силовыми линиями, моделирующих системы типа “Токамак”. Винтовая неустойчивость является одним из наиболее опасных видов МГД-неустойчивости для плазменных систем такого типа. Ее исследованию посвящено большое количество работ. В ранних работах устойчивость плазмы изучалась в линейном приближении аналитическими методами (см. работы [1–3] и библиографию к ним). Это давало только “опасные” значения параметров равновесной конфигурации, но не позволяло проследить ее дальнейшую эволюцию. Нелинейное рассмотрение задачи было впервые проведено Б.Б. Кадомцевым и О.П. Погуце на основе энергетических соображений [4]. Им принадлежит идея о проникновении в плазму вакуумных пузырей. В последние годы у нас в стране и за рубежом эволюцию неустойчивой плазмы успешно изучают, численно интегрируя систему МГД-уравнений на ЭВМ [5–25]. Совместная работа ИПМ и ИАЭ в этом направлении была начата по инициативе Б.Б. Кадомцева и нашла отражение в публикациях [7, 8, 10–13, 20, 22–24].

Данная работа продолжает наши предыдущие исследования. В ней подробно изучается зависимость инкремента винтовой неустойчивости от номера моды

<sup>\*)</sup> Двумерные численные модели плазмы. — М.: ИПМ, 1979. — С. 67–119 (совм. с Н.М. Зуевой, В.В. Палейчик).

при различных распределениях тока и детально анализируется поведение плазмы на нелинейной стадии при взаимодействии нескольких мод.

Изучение винтовой МГД-неустойчивости проводится в рамках модели “математического” тора, т.е. цилиндра с отождествленными торцами. Начальное равновесное состояние плазмы берется цилиндрически симметричным, а начальное возмущение — винтовое, т.е. зависящим в цилиндрической системе координат  $(r, \varphi, z)$  лишь от двух пространственных переменных: расстояния до оси  $r$  и винтовой переменной  $\theta = \varphi - \alpha z$ , где  $\alpha = 2\pi/L$ ,  $L$  — шаг винта. Радиус цилиндра принимается за единицу. В процессе эволюции винтовая симметрия плазмы сохраняется. Двумерность течения существенно облегчает возможность его расчета; в то же время модель довольно точно передает некоторые особенности “спиральных” движений в тороидальной камере.

В линейном приближении динамика винтового возмущения определяется модами  $\vec{V}_{mn}$  с

$$V_r = \frac{1}{r}(a_{mn} \cos m\theta + b_{mn} \sin m\theta) f_{mn}(r) \exp(\omega_{mn} t),$$

$m = 1, 2, \dots, n = 1, 2, \dots$ , где радиальный множитель  $f_{mn}(r)$  является собственной функцией дифференциального оператора (см. [26, формула (3.19)])

$$\left( \frac{rs}{1 - \varkappa} \frac{f'}{\beta} \right)' - \left\{ \frac{m^2 s}{r} - \frac{4\alpha^2 y^2 \beta_\varphi^2 r}{\beta s} + \left( \frac{1 + \varkappa b_\varphi^2}{1 - \varkappa} \right)' + \frac{\varkappa}{1 - \varkappa} \frac{\beta}{rs} (q + b_\varphi^2)^2 \right\} f = 0$$

с однородными граничными условиями.

Как показывают расчеты, наиболее опасны основные моды со знакопостоянным радиальным множителем  $f_{m0}$ ,  $f_{m0}(r) > 0$ ,  $0 < r < 1$ , развивающиеся каждая с существенно большим инкрементом, чем остальные моды серии  $f_{mn}$ . Последние с ростом номера локализуются в окрестности резонансной магнитной поверхности, на которой шаг винта силовых линий совпадает с  $L$  [23]. Однако первые моды возникают в порядке номеров при несколько более слабых токах, чем те, при которых в конфигурации возникает резонанс. Таким образом, область неустойчивых конфигураций оказывается шире чем вытекает из критерия Сайдема [27].

Для  $m = 1, 2, 3, 4$  было проведено сравнение инкрементов  $|\omega_m|$  основных мод при резонансе на магнитной оси для разных распределений тока. При параболическом токе ( $N = 1$ ) оказалось, что эти первые четыре моды как при  $\alpha = 1$ , так и при  $\alpha = 0,2$  имеют примерно равные инкременты. С приближением тока к однородному инкременты  $\omega_m(N)$  растут, стремясь к некоторому предельному значению  $\omega_m(\infty)$ . При этом большим номерам  $m$  мод соответствуют большие  $\omega_m(\infty)$  и  $\lim_{m \rightarrow \infty} \omega_m(\infty) < \infty$ .

Построение соответствующих собственных функций  $f_{m0}(r)$  позволяет сделать вывод, что при резонансе на оси моды практически локализованы в некоторой ее окрестности и быстро спадают вне ее. Для первых четырех мод области локализации практически совпадают, но при сравнении больших номеров  $m \sim 20$  эта область заметно стягивается. Такая локализация объясняет возникновение “круга неустойчивости”, проявляющегося на нелинейной стадии эволюции возмущений.

За развитием неустойчивости на нелинейной стадии удобно следить по эволюции магнитных поверхностей, точнее, по эволюции какого-либо нормального сечения семейства магнитных поверхностей. Существование таких поверхностей, на которые намотаны силовые линии, обеспечивается винтовой симметрией. В начальный момент они образуют систему коаксиальных круговых цилиндров, окружающих ось, что дает в нормальном сечении систему concentric окружностей.

Как было отмечено выше, когда резонансная силовая линия близка к оси цилиндра (или еще не появилась), основные моды  $f_{m0}$  быстро спадают к периферии. Тогда движение возмущенной плазмы происходит, как было показано в [12, 22, 24], практически лишь внутри некоторого цилиндра, т.е. в нормальном сечении — внутри “круга неустойчивости”. Внутри указанного цилиндра магнитные поверхности испытывают сильные деформации. Рассчитанная [12, 22, 24] картина показывает, что внутренние слои плазмы эффективно меняются местами с внешними (с сохранением топологии магнитных поверхностей), в результате чего образуются характерные узкие инверсные слои с большими градиентами плотности и магнитного поля, и, тем самым, большой плотностью электрического тока.

При увеличении равновесного тока резонансная поверхность отходит дальше от оси, а граница области неустойчивости быстро выходит на стенку цилиндра. В таких конфигурациях плазма из центра быстро выплескивается на кожух. Заметим, что основное дополнительное явление, к которому приводит учет конечной проводимости, заключается в расщеплении резонансной поверхности и образовании волокнистой структуры [13]; см. также [14, 16, 21, 22].

Проведенные расчеты нелинейной стадии эволюции для первых четырех мод и их линейных комбинаций показывают, что “круг неустойчивости” практически от  $m$  не зависит. Если каждая развивающаяся отдельная мода проявляет свои характерные особенности, то их комбинации, как и следовало ожидать, отличаются более сложным поведением. При этом черты одной моды могут нести одни магнитные поверхности, а черты другой — проявляться на других. В частности, на меньших радиусах очертания сечений могут быть близки к третьей, а на больших — к четвертой моде. Отметим, что сходная картина отмечалась при экспериментальном зондировании плазмы [28].

В качестве возмущающей заправки в начальный момент частицам плазмы придавалась скорость с гидродинамической функцией тока  $\Psi(r, \theta)$ :

$$\Psi(r, \theta) = \sum_{m=1}^s \lambda_m \frac{r^m}{m} (1 - r^2) \sin m\theta.$$

Компоненты такой скорости отличаются от собственных функций линейной МГД-задачи, хотя в наших расчетах близки к ним. Поэтому в наблюдаемой эволюции можно выделить этап выхода на собственное решение линейной задачи. После этого плазма начинает двигаться вслед за магнитными поверхностями (скорость плазмы направлена в основном по нормали к магнитным поверхностям). Однако, когда в ходе деформации поверхностей возникают узкие слои с большими градиентами магнитного поля и давления, поле начинает заметно

ускорять частицы плазмы, и они приобретают в этих местах большие скорости по касательной к магнитным поверхностям.

Следует отметить, что описанный выше нелинейный процесс развития неустойчивости носит достаточно общий характер и аналогично проявляется в ряде гидродинамических и магнитогидродинамических задач со “субстанциональным инвариантом”, т.е. когда существует некоторая физическая величина  $N$ , сохраняющая в каждой частице вещества свое первоначальное значение,  $dN/dT = 0$ . В качестве такого инварианта, в зависимости от рассматриваемого явления, может выступать энтропия, вращательный момент  $I = rv_\varphi$ , винтовая компонента скорости  $v_{\parallel} = v_z + drv_\varphi$ , величина  $B_\varphi/r\rho$  и в том числе рассматриваемая в задаче о винтовой МГД-неустойчивости функция потока  $\psi$ . При развитии неустойчивости поверхности уровня инварианта эволюционируют аналогично магнитным поверхностям, т.е. поверхностям уровня функции  $\psi$  [29, 30].

Авторы приносят глубокую благодарность за интерес к работе и полезные дискуссии Н.И. Герлах, К.В. Брушлинскому, С.В. Мирнову и Л.С. Соловьеву.

### § 1. Постановка задачи

В нашей математической модели рассматривается бесконечный плазменный цилиндр, находящийся во внешнем однородном продольном магнитном поле

$$B_z = 1 \quad (1.1)$$

и окруженный бесконечнопроводящим цилиндрическим кожухом радиуса  $r = 1$ . По плазме течет продольный ток  $j_z$  с плотностью

$$j_z = 2\mu_0 B_z (1 - r^{2N}). \quad (1.2)$$

Плазма предполагается изотермичной:

$$p = \mu_0^2 \rho, \quad (1.3)$$

и бесконечнопроводящей. Эти данные однозначно определяют соответствующую равновесную цилиндрически симметричную плазменную конфигурацию:

$$B_\varphi = \mu_0 r B_z \left(1 - \frac{r^{2N}}{N+1}\right), \quad (1.4)$$

$$\rho = B_z^2 \left[1 - r^2 - \frac{(N+2)(1 - r^{2N+2})}{(N+1)^2} + \frac{1 - r^{4N+2}}{(N+1)(2N+1)}\right], \quad (1.5)$$

$$B_r = 0, \quad j_r = 0, \quad j_\varphi = 0. \quad (1.6)$$

Согласно (1.6) компонента  $B_r$  магнитного поля равна нулю. Силовые линии такого поля описываются уравнениями

$$r = \text{const}, \quad \frac{d\varphi}{dz} = \frac{B_\varphi}{r B_z} = \mu_0 \left(1 - \frac{r^{2N}}{N+1}\right), \quad (1.7)$$

т.е. являются винтовыми линиями. Угол прокручивания каждой силовой линии зависит от  $r$  и параметров  $\mu_0$  и  $N$ :

$$\mu(r) = \mu_0 \left( 1 - \frac{r^{2N}}{N+1} \right), \quad (1.8)$$

Ему соответствует шаг винта

$$L_{\text{сл}} = \frac{2\pi}{\mu(r)}. \quad (1.9)$$

Коаксиальные цилиндры  $r = \text{const}$ , на которые навиваются силовые линии, являются магнитными поверхностями.

Таким образом, мы имеем семейство конфигураций, зависящее от параметра  $\mu_0$ , задающего силу тока, и параметра  $N$ , описывающего его распределение.

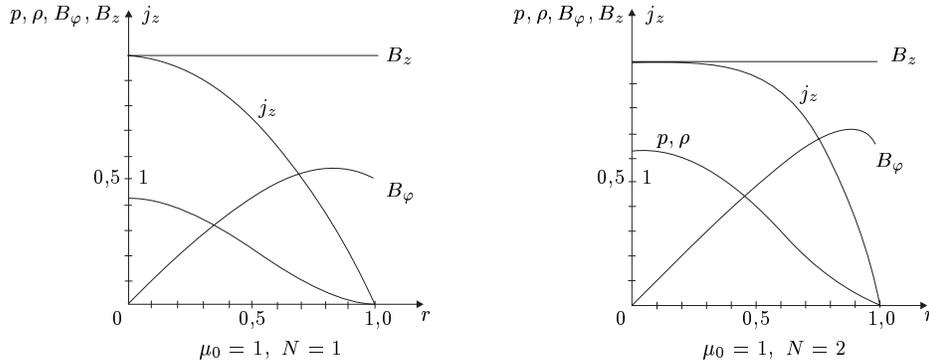


Рис. 1

(Параметр  $B_z = 1$  сохранен в формулах из соображений размерности.) Подчеркнем, что все равновесные величины  $p, \rho, \vec{B}, \vec{j}$  являются функциями одного только  $r$ . Профили этих величин при  $\mu_0 = 1, N = 1, 2$  изображены на рис. 1.

Наша цель — изучить устойчивость плазменного шнура относительно винтовых возмущений, т.е. возмущений  $f(r, \varphi, z)$  вида

$$f(r, \varphi, z) = f(r, \theta), \quad (1.10)$$

$$\theta = \varphi - \frac{2\pi z}{L}, \quad (1.11)$$

в зависимости от  $\mu_0, N$  и шага  $L$  винта.

В первую очередь мы изучим устойчивость равновесной конфигурации относительно бесконечно малых винтовых возмущений. Как известно, наиболее опасны возмущения, постоянные вдоль силовых линий, т.е. когда шаг винта возмущения совпадает с шагом винта силовой линии. Грубо говоря, такие резонансные линии начинают перемещаться как целое. Разумеется, наличие резонанса само по себе не обязательно влечет неустойчивость. Расстройка резонанса на соседних магнитных поверхностях может явиться стабилизирующим фактором.

Как показал Сайдем [27], равновесная цилиндрическая конфигурация заведомо неустойчива, когда на резонансной поверхности

$$\left(\frac{d\mu}{dr}\right)^2 r^2 B_z \leq -4\mu^2 \frac{dp}{dr}, \quad (1.12)$$

т.е. резонанс достаточно широк. Обратное неравенство дает необходимый критерий устойчивости

$$\frac{\mu'^2 r^2}{4} + \frac{\mu^2 p'}{B_z} > 0 \quad (1.13)$$

в терминах градиентов плотности  $p' = dp/dr$  и угла прокручивания  $\mu' = d\mu/dr$ .

В своей работе Сайдем указал специальный тип возмущений, локализованных около резонансной поверхности и экспоненциально возраставших, когда выполнено условие (1.12). Такой подход выделяет достаточно широкий класс неустойчивых конфигураций, но не позволяет точно определить область параметров, отвечающих устойчивым конфигурациям. Нами использован спектральный подход. Рассматривая для конкретной конфигурации линеаризованную систему уравнений для бесконечно малых возмущений, мы вычисляем и исследуем собственные функции. Таким образом нам удается точно установить, устойчива или нет взятая конфигурация относительно винтовых возмущений. Далее, когда конфигурация неустойчива, мы интегрируем систему МГД-уравнений, прослеживаем эволюцию неустойчивой моды или комбинации таких мод (точнее, эволюцию возмущений, близких к собственным модам).

Вместо бесконечного цилиндра рассмотрим теперь цилиндр конечной длины  $l$  с отождествленными торцами. Для нас он будет удобной математической моделью тора с длиной  $l$  магнитной оси (и длиной меридиана  $2\pi$ ). Движения плазмы внутри такого цилиндра совпадают с периодическими движениями плазмы внутри бесконечного цилиндра, имеющими период  $l$ .

Условие периодичности ограничивает класс допустимых винтовых возмущений  $f(r, \theta)$ , поскольку необходимо

$$f\left(r, \varphi - \frac{2\pi l}{L}\right) \equiv f(r, \varphi), \quad (1.14)$$

где  $L$  — шаг винта. Это соотношение выполняется автоматически для любой функции  $f$  при  $l = kL$ , выполняется только для  $f(r, \theta) \equiv f(r)$  при иррациональных  $l/L$  и справедливо лишь для периодических функций

$$f(r, \varphi) \equiv f\left(r, \varphi + \frac{2\pi}{k}\right) \quad (1.15)$$

при рациональном отношении  $l/L = \nu/k$ , где дробь  $\nu/k$  несократима.

Ввиду периодичности рассматриваемого цилиндра по  $z$  и по  $\varphi$  все возмущения  $f(r, \varphi, z)$  разложимы в двойной ряд Фурье:

$$f(r, \varphi, z) \sim \sum f_{m,n}(r) \exp\left(im\varphi + 2\pi in\frac{z}{l}\right), \quad (1.16)$$

т.е. все собственные вектор-функции являются винтовыми,  $m\varphi + 2\pi n z/l = m(\varphi - \alpha z)$ ,  $\alpha = -2\pi n/(ml)$ . Поэтому устойчивость конечного периодического цилиндра относительно винтовых возмущений (включая аксиально симметричные при  $m = 0$ ) влечет его устойчивость относительно любых бесконечно

малых возмущений. (С некоторыми оговорками это же справедливо и для бесконечного цилиндра.) Бесконечный плазменный цилиндр со спадающей к кожуху плотностью,  $p' < 0$ , очевидным образом неустойчив относительно винтовых возмущений, поскольку в малой окрестности магнитной оси для каждой магнитной поверхности существует неустойчивое резонансное возмущение Сайдема. Поэтому для бесконечного цилиндра можно говорить только об устойчивости или неустойчивости отдельных мод и их инкрементах. Для конечного же цилиндра ввиду требования периодичности (1.15) возмущений такой тривиальный вывод уже необоснован. Так как периодический цилиндр физически нереализуем, мы не ставили себе задачу построить до конца теорию его винтовой неустойчивости. В первую очередь нас интересовали качественные результаты, которые в какой-то мере справедливы и для физических торов. Вопрос же об устойчивости относительно всех мод оставлен нами открытым.

## § 2. Методика расчета эволюции конечных винтовых возмущений

Предполагается, что в начальный момент частицы плазмы испытывают “удар” и приобретают некоторые скорости  $\vec{v}(r, \theta)$ ,

$$\theta = \varphi - \alpha z, \quad \alpha = \frac{2\pi}{L}, \quad (2.1)$$

где  $L$  — шаг винта возмущений. Эти скорости задаются гидродинамической функцией тока вида

$$\Psi(r, \theta) = \sum_{m=1}^s \lambda_m \frac{r^m}{m} (1 - r^2) \sin m\theta. \quad (2.2)$$

Здесь  $\lambda_m$  — амплитуды возмущений, меняющиеся от варианта к варианту; во всех расчетах было  $s \leq 4$ . При таком задании поле начальных скоростей удовлетворяет условию  $\text{div} \vec{v} = 0$ . Сами компоненты скорости принимались равными соответственно

$$v_r = \frac{1}{r} \frac{\partial \Psi}{\partial \theta} = \sum_{m=1}^s \lambda_m r^{m-1} (1 - r^2) \cos m\theta, \quad (2.3)$$

$$v_\varphi = -\frac{\partial \Psi}{\partial r} = -\sum_{m=1}^s \lambda_m r^{m-1} \left[ 1 - \left( 1 + \frac{2}{m} \right) r^2 \right] \sin m\theta, \quad (2.4)$$

$$v_z = 0. \quad (2.5)$$

Магнитное поле в момент удара остается невозмущенным, т.е. описывается формулами (1.1), (1.4), (1.6), но возникшее движение плазмы начинает его изменять.

Движение плазмы и изменение магнитного поля мы описываем уравнениями

идеальной магнитной гидродинамики

$$\frac{\partial p}{\partial z} + \operatorname{div} \rho \vec{v} = 0, \quad (2.6)$$

$$\rho \frac{d\vec{v}}{dt} = -\nabla p + [\vec{j}, \vec{B}], \quad (2.7)$$

$$\frac{\partial \vec{B}}{\partial t} = \operatorname{rot} [\vec{v}, \vec{B}], \quad (2.8)$$

$$\frac{\partial p}{\partial t} + (\vec{v}, \nabla)p + \gamma p \operatorname{div} \vec{v} = 0, \quad (2.9)$$

где  $\rho$  — плотность,  $p$  — давление,  $\vec{v}$  — скорость,  $\vec{B}$  — напряженность магнитного поля,  $\vec{j} = \operatorname{rot} \vec{B}$  — плотность электрического тока. В расчетах всегда бралось  $\gamma = 1$ , т.е. рассматривалась изотермичная плазма. Предполагается, что эволюция винтовых возмущений сохраняет винтовую симметрию задачи. Поэтому уравнения (2.6)–(2.9) записываются в винтовой системе координат  $(r, \theta)$  и в согласии с (2.1) принимается

$$f(r, \varphi, z) = f(r, \varphi - \alpha z), \quad \frac{\partial}{\partial \varphi} = \frac{\partial}{\partial \theta}, \quad \frac{\partial}{\partial z} = -\alpha \frac{\partial}{\partial \theta}. \quad (2.10)$$

Полученная система численно интегрируется разностными методами в области

$$0 \leq r \leq 1, \quad 0 \leq \theta \leq 2\pi. \quad (2.11)$$

На кожухе  $r = 1$  ставятся граничные условия

$$v_r = 0, \quad B_r = 0, \quad E_\tau = 0. \quad (2.12)$$

Условие  $E_\tau = 0$  приводит к  $\partial B_z / \partial r = 0$ ,  $\partial r B_\varphi / \partial r = 0$ .

На оси  $r = 0$  особенность имеет только система цилиндрических координат, а все скалярные и векторные поля должны быть гладкими. Это требование дает граничные условия для численного интегрирования. Значения всех величин при  $\theta = 0$  и  $\theta = 2\pi$  связаны условием периодичности. Практически при расчете отдельно четных и нечеткие возмущений можно сократить объем работы ЭВМ, уменьшая область интегрирования до  $0 \leq \theta \leq \pi$ ,  $0 \leq r \leq 1$ .

Естественный класс возмущений получается при

$$\Psi = \sum_{m=1}^s \lambda_m \frac{r^m}{m} (1 - r^2) (a_m \sin m\theta + b_m \cos m\theta).$$

Для экономии мы ограничились более узким классом (2.2), что также позволило сократить область интегрирования до  $0 \leq \theta \leq \pi$  ввиду условия антисимметрии для  $\Psi$  как при  $\theta = 0$ , так и  $\theta = \pi$ .

Эта антисимметрия очень упрощает граничные условия при  $r = 0$ :

$$v_r(0, \theta) = v_r(0, 0) \cos m\theta, \quad v_\varphi(0, \theta) = -v_r(0, 0) \sin m\theta, \quad v_z(0, \theta) = 0. \quad (2.13)$$

Последнее равенство следует из условия  $v_z(r, 0) \equiv 0$ , вытекающего из антисимметричности, и условия  $\partial v_z / \partial \theta(0, \theta) = 0$ , вытекающего из гладкости поля. Сходным образом

$$B_r(0, \theta) = B_\varphi(0, 0) \sin m\theta, \quad B_\varphi(0, \theta) = B_\varphi(0, 0) \cos m\theta, \quad B_z(0, \theta) = B_z(0, 0). \quad (2.14)$$

За движением плазмы при развитии винтовой неустойчивости удобно следить по линиям уровня функции потока  $\psi$ . Функция  $\psi$ , удовлетворяющая системе

$$\frac{\partial \psi}{\partial r} = -(B_\varphi - \alpha r B_z), \quad \frac{\partial \psi}{\partial \theta} = r B_r, \quad (2.15)$$

является первым интегралом для системы уравнений силовых линий. Для идеальной (бесконечнопроводящей) плазмы для каждой частицы плазмы она сохраняет свое начальное значение:

$$\frac{d\psi}{dt} = \frac{\partial \psi}{\partial t} + v_r \frac{\partial \psi}{\partial r} + \left( \frac{v_\varphi}{r} - \alpha v_z \right) \frac{\partial \psi}{\partial \theta} = 0. \quad (2.16)$$

Поэтому в начальный момент мы выбираем систему точек, лежащих на избранных магнитных поверхностях, и численно интегрируем уравнения движения этих точек

$$\frac{dr}{dt} = v_r, \quad \frac{d\theta}{dt} = \frac{v_\varphi}{r} - \alpha v_z \quad (2.17)$$

на картинной плоскости  $(r, \theta)$  по уже вычисленному разностными методами полю скоростей. В результате мы получаем систему опорных точек, описывающих структуру магнитных поверхностей в текущий момент времени.

В начальный момент магнитные поверхности образуют систему коаксиальных цилиндров. Соответственно на картинной плоскости  $(r, \theta)$  мы видим систему концентрических окружностей. Величина  $\psi$  в начальный момент легко определяется из (2.15) интегрированием:

$$\psi(r, \theta) = \frac{(\alpha - \mu_0)r^2}{2} + \frac{\mu_0 B_z r^{2N+2}}{2(N+1)^2}. \quad (2.18)$$

В ходе эволюции первоначальная картина изменяется весьма существенно.

Функцию потока  $\psi$  можно было бы считать и иначе, непосредственно приближенно интегрируя в каждый счетный момент времени систему (2.15) с уже вычисленными правыми частями или же численно интегрируя уравнение (2.16), определяющее  $\partial \psi / \partial t$  в каждой точке эйлеровой разностной сетки. Однако обоим этим способам присущ общий недостаток — они искажают топологию системы магнитных поверхностей (точнее, их следов на картинной плоскости  $(r, \theta)$ ). Из-за присущего разностным схемам сглаживания приближенные значения функции  $\psi$  в окрестности минимума начинают увеличиваться и притом в разных узлах сетки по-разному. Возникает чисто счетный эффект “якобы расщепления” магнитных поверхностей, который противоречит вмерзности силовых линий в идеальную плазму. Если в начале эволюции путем дополнительной обработки еще удастся отделить реальную физическую картину процесса от чисто счетных эффектов, то в ходе дальнейшего развития возмущений это становится

невозможным. Рассчитывая же траектории движения частиц, мы лишь несколько ошибаемся в положении магнитных поверхностей, а сами значения функции  $\psi$  сохраняются постоянными. Заметим еще, что благодаря “инерционности” расчета траекторий наш прием дает достаточно точное изображение конфигурации, когда рассчитываемые магнитные поля уже сильно “смазались” из-за неизбежной схемной “вязкости”. Подобный прием оказывается удобным и при расчете конечно-проводящей плазмы. Только здесь значение  $\psi$  для частиц меняется со временем, и его приходится досчитывать дополнительным интегрированием по траектории частицы.

### § 3. Методика расчета винтовой МГД-неустойчивости в линейном приближении

Пусть величины  $p_0(r, \varphi, z)$ ,  $\vec{B}_0(r, \varphi, z)$ ,  $\rho_0(r, \varphi, z)$  описывают равновесное состояние идеальной плазмы, т.е.

$$\vec{v}_0 = 0, \quad \nabla p_0 = [\text{rot } \vec{B}_0, \vec{B}_0]. \quad (3.1)$$

Эти величины подвергаются малому возмущению:

$$\begin{aligned} \rho &= \rho_0 + \rho_1(t; r, \varphi, z), & p &= p_0 + p_1(t; r, \varphi, z), \\ \vec{B} &= \vec{B}_0 + \vec{B}_1(t; r, \varphi, z), & \vec{v} &= \vec{v}_0 + \vec{v}_1(t; r, \varphi, z). \end{aligned} \quad (3.2)$$

Возмущенные величины должны удовлетворять МГД-уравнениям (2.6)–(2.9). Пренебрегая в них слагаемыми второго порядка малости, получаем линеаризованную систему МГД-уравнений вблизи равновесия

$$\frac{\partial p_1}{\partial t} = -\rho_0 \text{div } \vec{v}_1 - \vec{v}_1 \nabla p_0, \quad (3.3)$$

$$\frac{\partial p_1}{\partial t} = -\vec{v}_1 \nabla p_0 - \gamma p_0 \text{div } \vec{v}_1, \quad (3.4)$$

$$\frac{\partial \vec{B}_1}{\partial t} \text{rot } [\vec{v}_1, \vec{B}_0], \quad (3.5)$$

$$\rho_0 \frac{\partial \vec{v}_1}{\partial t} = -\nabla p_1 + [\text{rot } \vec{B}_0, \vec{B}_1] + [\text{rot } \vec{B}_1, \vec{B}_0]. \quad (3.6)$$

Систему (3.3)–(3.6) можно свести [1] к одному векторному уравнению

$$\rho_0 \frac{\partial^2 \vec{v}_1}{\partial t^2} = \nabla (\vec{v}_1 \nabla p_0 + \gamma p_0 \text{div } \vec{v}_1) + [\text{rot } \vec{B}_0, \text{rot } [\vec{v}_1, \vec{B}_0]] + [\text{rot rot } [\vec{v}_1, \vec{B}_0], \vec{B}_0]. \quad (3.7)$$

Это уравнение для возмущения скорости вместе с граничными условиями, вытекающими из (2.11), (2.12), и соответствующими начальными условиями определяет эволюцию возмущений вблизи состояния равновесия. Его удобно записать в операторной форме

$$\rho_0 \frac{\partial^2 \vec{v}_1}{\partial t^2} = -\mathbb{K}[\vec{v}_1], \quad (3.8)$$

где линейный оператор  $\mathbb{K}$ , определяемый правой частью (3.7), является самосопряженным [1].

Как известно, поведение решений линейного уравнения (3.8) может быть изучено методом Фурье. Для этого надо найти решения, имеющие вид

$$\vec{v}_1(t; r, \varphi, z) = e^{-i\omega t} \vec{v}_2(r, \varphi, z). \quad (3.9)$$

Зависящий только от пространственных переменных множитель  $\vec{v}_2(r, \varphi, z)$  должен удовлетворять однородному уравнению

$$\mathbb{K}[\vec{v}_2] + \omega^2 \rho_0 \vec{v}_2 = 0, \quad (3.10)$$

т.е. быть собственной функцией оператора  $\mathbb{K}/\rho_0$ , отвечающей собственному числу  $\lambda = \omega^2$ .

Собственные значения  $\lambda = \omega^2$  самосопряженного оператора всегда действительны. Возможны два случая:

$$1^\circ) \lambda = \omega^2 \geq 0; \quad 2^\circ) \lambda = \omega^2 \leq 0. \quad (3.11)$$

Если  $\omega^2 \geq 0$ , то собственная мода устойчива и осциллирует по времени по закону  $e^{\mp i\sqrt{\lambda}t}$ . Если же  $\omega^2 < 0$ , то мода неустойчива и возрастает по закону  $e^{|\omega|t}$ .

Собственно говоря, каждой собственной вектор-функции  $\vec{v}_2(r, \varphi, z)$  отвечают два инкремента  $\pm\omega$ . Это объясняется тем, что система собственных функций  $\{\vec{v}_2\}$  является перепополненной в смысле М.В. Келдыша [31], так как определяет систему собственных вектор-функций вектор-оператора, определяемого всеми уравнениями (3.3)–(3.6). Поясним сказанное примером. Пусть малое возмущение задано собственным полем скоростей  $\vec{v}_2$ , а возмущения магнитного поля и давления в начальный момент равны нулю. Из первого условия имеем

$$\vec{v}_1(t) = [c_1 e^{-i\omega t} + c_2 e^{i\omega t}] \vec{v}_2, \quad c_1 + c_2 = 1.$$

Из второго условия вытекает, что в начальный момент правая часть (3.6) равна нулю. Отсюда

$$0 = \rho \frac{\partial \vec{v}_1}{\partial t} \Big|_{t=0} = \rho_0 [-i\omega c_1 + i\omega c_2].$$

Следовательно,  $c_1 = c_2 = 1/2$ , и эволюция указанного возмущения описывается полем скоростей

$$\begin{aligned} \vec{v}_1(t) &= \vec{v}_2 \cos \omega t \quad \text{при} \quad \omega^2 \geq 0, \\ \vec{v}_1 &= \vec{v}_2 \operatorname{ch} |\omega|t \quad \text{при} \quad \omega^2 < 0. \end{aligned} \quad (3.12)$$

Остальные величины могут быть найдены интегрированием (3.3)–(3.5). Принципиальной роли это замечание не играет; важно лишь, что появление собственного числа  $\lambda = \omega^2 < 0$  влечет неустойчивость соответствующей моды.

В нашей задаче исходное равновесное состояние является цилиндрически симметричным, т.е.  $p_0, \rho_0, \vec{B}_0$  зависят только от  $r$ ,  $\vec{v}_0 = 0$ ,  $B_{r_0} = 0$ , а соотношение (3.1) между  $p_0$  и  $\vec{B}_0$  упрощается до

$$\frac{\partial p_0}{\partial r} - B_{z_0} \frac{\partial B_{z_0}}{\partial r} - \frac{B_{\varphi_0}}{r} \frac{\partial (r B_{\varphi_0})}{\partial r}. \quad (3.13)$$

На это равновесное состояние накладывается бесконечно малое собственное винтовое возмущение, которое задается возмущением скорости

$$\vec{v}_1(t; r, \varphi, z) = e^{-i\omega t} \vec{v}(r, \theta), \quad (3.14)$$

где  $\theta = \varphi - \alpha z$ ,  $\alpha = 2\pi/L$ ,  $L$  — шаг винта; см. (2.1).

Для компонент скорости, зависящих только от  $r$  и  $\theta$ , операторы в (3.7) и (3.10) по правилу (2.10) переходят в дифференциальные операторы также только от  $r$  и  $\theta$ , инвариантные относительно группы вращений по углу  $\theta$ . Поэтому возможно дальнейшее разделение переменных:

$$v_r(r, \theta) = v_{r,m}(r) \cos m\theta, \quad (3.15)$$

$$v_\varphi(r, \theta) = v_{\varphi,m}(r) \sin m\theta, \quad (3.16)$$

$$v_z(r, \theta) = v_{z,m}(r) \sin m\theta. \quad (3.17)$$

Вместо  $\theta$  в тригонометрические функции можно подставить  $\theta - \theta_0$  при любом  $\theta_0$ , так что имеется ровно два линейно независимых собственных вектор-функции  $\vec{v}_m(r, \varphi)$ , отвечающих вектор функции  $\vec{v}_m(r)$ , но для исследования устойчивости это роли не играет.

Уравнение (3.7) для собственной винтовой вектор-функции

$$\vec{v}_2(r, \theta) = \vec{v}_m(r) \begin{cases} \cos m\theta \\ \sin m\theta \end{cases}$$

удаётся свести к уравнению для одной скалярной функции

$$f(r) = r v_r(r), \quad (3.18)$$

где мы опускаем индексы  $m$  и  $\alpha$ . Такое уравнение в общем случае стационарного невозмущенного течения было получено Л.С. Соловьевым [26]. В нашем частном случае, когда  $\vec{v}_0 \equiv 0$ , оно сводится к

$$\frac{d}{dr} \left( \frac{rs}{(1-\varkappa)\beta} \frac{df}{dr} \right) - \left\{ \frac{m^2 s}{r} - \frac{4\alpha^2 y^2 b_{\varphi_0}^2 r}{\beta s} + \frac{d}{dr} \left( \frac{q + \varkappa b_{\varphi_0}^2}{1-\varkappa} \right) + \frac{\varkappa}{1-\varkappa} \frac{\beta}{rs} (q^2 + b_{\varphi_0}^2)^2 \right\} f = 0, \quad (3.19)$$

где введены сокращения

$$\begin{aligned} \beta &= 1 + \alpha^2 r^2, & \beta_{\varphi_0} &= \frac{B_{\varphi_0}}{r}, & y &= \alpha B_{z_0} - b_{\varphi_0}, & B_0^2 &= B_{\varphi_0}^2 + B_{z_0}^2, \\ s &= \frac{\rho_0 \omega^2}{m^2} - y^2, & q &= \frac{2y b_{\varphi_0}}{\beta} + b_{\varphi_0}^2, & \varkappa &= r^2 \frac{\rho_0 \omega^2}{\beta m^2} \left( B_0^2 + \frac{m^2 s \gamma p_0}{\rho_0 \omega^2} \right)^{-1}. \end{aligned} \quad (3.20)$$

Нас интересует только случай  $\omega^2 < 0$ . Тогда  $s < 0$ ,  $-\infty < \varkappa < 0$ , т.е. уравнение (3.19) на отрезке  $[0, 1]$  имеет единственную особенность при  $r = 0$ . Поэтому однородные граничные условия имеют вид

$$f(r) \asymp r^m \quad \text{при} \quad r \rightarrow 0, \quad (3.21)$$

$$f(1) = 0. \quad (3.22)$$

Коэффициенты уравнения (3.19) зависят от равновесных значений  $\rho_0, p_0, \vec{B}_0$  от параметра  $\alpha$  винта и углового номера  $m$  моды, а также от параметра  $\lambda = \omega^2$  квадрата инкремента. Уравнение (3.19) принадлежит к уравнению типа Штурма–Лиувилля. Нас интересуют те значения  $\lambda$ , при которых уравнение имеет нетривиальное решение. Из теории операторов Штурма–Лиувилля и характера зависимости коэффициентов от  $\lambda$  вытекает, что существует последовательность  $f_0(r), f_1(r), f_2(r), \dots$  нетривиальных решений, отвечающих каким-то значениям  $\lambda_0, \lambda_1, \lambda_2, \dots$ . При этом  $f_0(r)$  не имеет нулей на интервале  $0 < r < 1$ ,  $f_1(r)$  имеет ровно один нуль,  $f_2(r)$  — ровно два нуля и т.д. Мода  $v_k(r) = f_k(r)/r$  неустойчива, если соответствующее  $\lambda_k < 0$ .

Собственные функции  $f_k(r)$  при заданных значениях параметра  $\alpha$  и  $m$  мы искали методом пристрелки. Для уравнения (3.19) при некотором  $\omega^2 = \lambda$  разностным методом решалась задача Коши с начальным условием  $f(r, \lambda) = r^m$  и определялась величина  $f(1, \lambda)$ . Затем параметр  $\lambda$  варьировался так, чтобы  $f(1, \lambda)$  обратилось в нуль. Тем самым определялось нетривиальное решение краевой задачи.

#### § 4. Результаты расчета винтовой МГД-неустойчивости в линейном приближении

Рассмотрены винтовые возмущения с двумя значениями параметра  $\alpha$ :  $\alpha_1 = 1$  и  $\alpha_2 = 1/5$ . Им соответствуют шаги винта  $L_1 = 2\pi$  и  $L_2 = 10\pi$ . Используя уравнение (3.19), для каждого из этих значений  $\alpha$  решались две задачи.

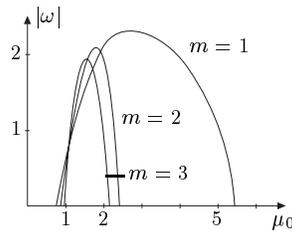
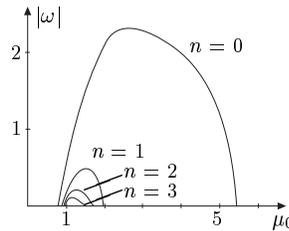
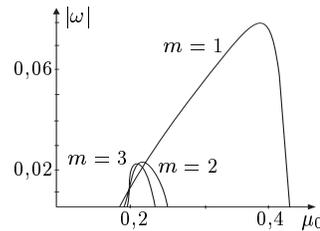
I. При заданных значениях параметров  $B_z = 1$  и  $N = 1$ ,  $m = 1, 2, 3$ , определялась область значений параметра  $\mu_0$ , при которых развивается винтовая неустойчивость, а также вычислялись инкременты этой неустойчивости.

II. При резонансе в центре, т.е. при  $\mu_0 = \alpha$  и при  $B_z = 1$ ,  $m = 1, 2, 3, 4$ , определялись инкременты винтовой неустойчивости в зависимости от  $N$ , т.е. в зависимости от различных профилей тока. При этом изучалась и форма радиальных собственных функций.

Кроме того, для  $\alpha = 1$ ,  $B_z = 1$ ,  $N = 1$ ,  $m = 1$  была найдена последовательность радиальных собственных функций соответственно с  $n = 0, 1, 2, 3$  нулями, а также отвечающие им инкременты.

Результаты расчетов приведены на рис. 2–7.

На рис. 2 и 4 представлены результаты, относящиеся к задаче I, а именно приведена зависимость инкремента  $|\omega|$  от  $\mu_0$  для  $\alpha = 1$  и  $\alpha = 1/5$  соответствен-

Рис 2.  $\alpha = 1$ Рис 3.  $\alpha = 1$ Рис 4.  $\alpha = 0,2$

но. Заметим, что значения  $\mu_0 = 1$  и  $\mu_0 = 2$  для  $\alpha = 1$  и  $\mu_0 = 0,2$  и  $\mu_0 = 0,4$  для  $\alpha = 1/5$  являются пределами существования резонансных поверхностей, так как при  $\mu_0 = 1$  и соответственно  $\mu_0 = 0,2$  радиус резонансной поверхности  $r_s = 0$ , а при  $\mu_0 = 2$  и  $\mu_0 = 0,4$  —  $r_s = 1$ . Из рис. 2 следует, что для больших  $\alpha$  граница области неустойчивости при  $m = 1, 2, 3$  с обеих сторон выходит за пределы существования резонансных поверхностей. Это утверждение справедливо для первой моды  $m = 1$  и в случае  $\alpha = 0,2$  (см. рис. 4). В то же время вторая и третья моды, несмотря на наличие резонанса, устойчивы при  $\mu_0 > 0,25$ . Следует отметить, что в этом случае выполняется необходимый критерий устойчивости Сайдема (1.13); ср. [2, 23].

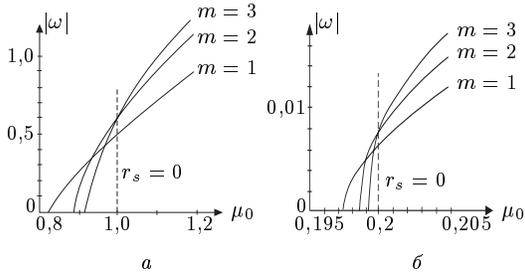


Рис 5. а:  $\alpha = 1$ ; б:  $\alpha = 0,2$

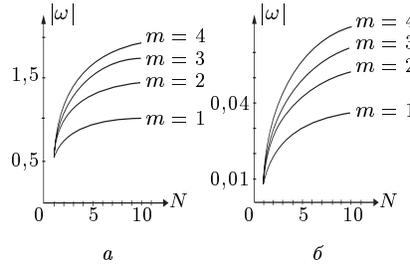


Рис. 6.  $j_z = 2\mu_0(1 - r^{2N})$ . а:  $\alpha = 1$ ,  $\mu_0 = 1$ ; б:  $\alpha = 0,2$ ,  $\mu_0 = 0,2$

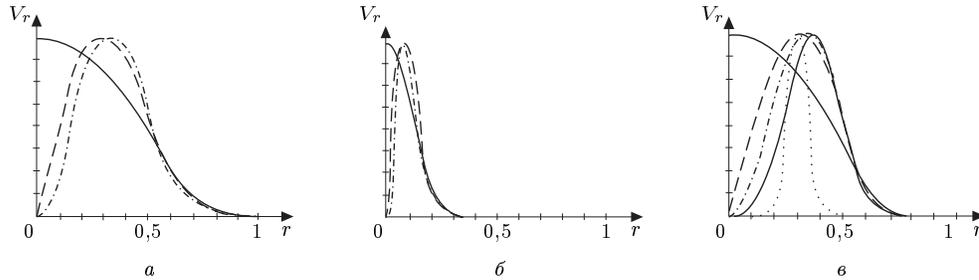


Рис. 7. а:  $\alpha = 1$ ,  $\mu_0 = 1$ ,  $N = 1$ ; б:  $\alpha = 0,2$ ,  $\mu_0 = 0,2$ ,  $N = 1$ ; в:  $\alpha = 0,2$ ,  $\mu_0 = 0,2$ ,  $N = 2$

На рис. 5, а изображена в увеличенном масштабе в окрестности  $\mu_0 = 1$  и  $\mu_0 = 0,2$  соответственно та же зависимость инкремента  $|\omega|$  от  $\mu_0$ , что на рис. 2 и рис. 4.

Из рис. 5, а видно, что:

- 1) неустойчивость возникает раньше (т.е. при  $\mu_0 < \alpha$ ), чем наступает резонанс в нуле (когда  $\mu_0 = \alpha$ ), о чем уже говорилось выше;
- 2) прежде всего неустойчивой становится первая мода  $m = 1$ ;
- 3) при  $\mu_0 = 1$  рассматриваемые моды  $m = 1, 2, 3$  имеют примерно одинаковые инкременты винтовой неустойчивости, причем для  $m = 2$  и для  $m = 3$  они даже немного больше, чем для  $m = 1$ .

Из рис. 5, б следует, что аналогичная закономерность справедлива и для малых  $\alpha$ .

На рис. 3<sup>1)</sup> приведена зависимость  $|\omega|$  от  $N$  для радиальных собственных функций с  $n$  нулями при  $m = 1$ ,  $\alpha = 1$ ,  $B_z = 1$ ,  $N = 1$ ,  $n = 0, 1, 2, 3$ . Видно, что инкременты падают при увеличении числа нулей, а размеры области неустойчивости сокращаются. Неустойчивой становится сначала основная мода без нулей.

На рис. 6, *a*, *б* изображена зависимость  $|\omega|$  от  $N$  в случае резонанса в центре. Значение  $N = 1$  соответствует параболическому току. При увеличении  $N$  происходит уплощение тока в центре, так что он все больше стремится к однородному,  $\mu = \mu_0(1 - r^{2N}/(N + 1)) \rightarrow \mu_0$ .

Для мод с номерами  $m = 1, 2, 3, 4$  при  $\mu_0 = \alpha$  справедливо следующее утверждение: моды с большими номерами имеют большие инкременты. Исключение составляет при  $N = 1$  сравнение мод  $m = 3$  и  $m = 4$ .

Поведение инкремента с ростом номера  $m$  моды при больших  $N$  и малых  $\alpha$  можно описать формулой из [23]

$$\frac{\rho\omega_{mn}^2}{m^2} = -\frac{4\alpha^4}{x_{mn}^2},$$

где  $x_{mn}$  — корни бесселевой функции  $J_m(x)$ .

При больших  $m$  первый корень бесселевой функции  $J_m(x)$ , как известно, имеет следующую асимптотику [32]:  $x_m = m + m^{1/3} \cdot 1,8557 + O(m^{-1/3})$ , откуда

$$\begin{aligned} |\omega_m| &= \\ &= \frac{2\alpha^2 m}{\sqrt{\rho}(m + m^{1/3} 1,8557 + O(m^{-1/3}))} = \frac{2\alpha^2}{\sqrt{\rho}(1 + 1,8557(m^{1/3}/m) + O(m^{-1/3}/m))}. \end{aligned}$$

Таким образом, при увеличении  $m$  инкременты увеличиваются, стремясь к пределу  $2\alpha^2/\sqrt{\rho}$ . При  $\alpha = 0,2$  и  $\rho = 1$   $|\omega| \rightarrow 0,08$ , что согласуется с нашими результатами  $|\omega_4| \approx 0,007$ .

На рис. 7 изображены радиальные собственные функции, причем:

на рис. 7, *a* для  $\alpha = 1$   $B_z = 1$ ,  $N = 1$ ,  $\mu_0 = 1$ ,  $m = 1, 2, 3$ ;

на рис. 7, *б* для  $\alpha = 1/5$   $B_z = 1$ ,  $N = 1$ ,  $\mu_0 = 1/5$ ,  $m = 1, 2, 3$ ;

на рис. 7, *в* для  $\alpha = 1/5$   $B_z = 1$ ,  $N = 2$ ,  $\mu_0 = 1/5$ ,  $m = 1, 2, 3, 4, 20$ .

Из сравнения рис. 7, *a* и 7, *б* следует, что при малых  $\alpha$  собственные функции сильно локализованы в окрестности нуля, причем для рассмотренных  $m = 1, 2, 3$  область локализации примерно одинаковая. Результаты, приведенные на рис. 7, *б* и 7, *в*, показывают роль  $N$ . При увеличении  $N$  не только увеличиваются инкременты, но и увеличивается область локализации собственной функции. Рис. 7, *в* позволяет увидеть, как изменяется радиальная собственная функция  $v_r$  с ростом  $m$ . При  $m = 10$  собственная функция заметно обострена (на чертеже ее график не изображен), при  $m = 20$  она обострена еще больше, а область ее локализации передвинута к точке резонанса  $r = 0$ .

Таким образом, спектральный подход позволил установить, что неустойчивость первых винтовых мод имеет место и в случае отсутствия резонанса. Достаточно существования силовых линий с шагом винта, близким к шагу винта возмущений. И чем система ближе к резонансу, тем больше неустойчивых мод и тем больше их инкременты.

<sup>1)</sup>Расчеты, представленные на рис. 2–5, выполнены Н.И. Герлах.

**§ 5. Результаты расчета винтовой МГД-неустойчивости в нелинейном случае**

В предыдущем параграфе рассматривалась винтовая неустойчивость в линейном приближении. В частности, там было установлено (см. рис. 6), что при резонансе в центре в случае параболического тока ( $N = 1$ ) инкременты винтовой неустойчивости при  $m = 1, 2, 3, 4$  почти равны, а при распределении тока с  $N > 1$  возрастают вместе с номером  $m$ . Это справедливо как при больших  $\alpha$  ( $\alpha = 1$ ), так и при малых ( $\alpha = 0,2$ ). Именно в связи с этим возник вопрос об изучении динамики развития винтовой неустойчивости в случае взаимодействия нескольких мод.

Для этого решалась сформулированная в § 1 задача в нелинейном случае. В качестве начальных параметров брались (см. рис. 1)

$$\alpha = 1, \quad \mu_0 = 1, \quad B_z = 1, \quad N = 1, 2, 6.$$

Возмущение скорости определялось по формулам (2.3)–(2.5):

$$v_r = \sum_{m=1}^s \lambda_m r^{m-1} (1 - r^2) \cos m\theta, \quad v_\varphi = - \sum_{m=1}^s \lambda_m r^{m-1} \left[ 1 - \left( 1 + \frac{2}{m} \right) r^2 \right] \sin m\theta, \\ v_z = 0.$$

Отметим, что естественнее было бы брать в качестве компонент возмущения скорости не произвольные функции, а собственные, которые определяются из линейной теории. Как показывают расчеты, если затравочные скорости сильно отличаются от собственных функций, то значительно удлиняется время решения задачи. Счет происходит так, что сначала вырабатывается растущая собственная вектор-функция, а затем начинается собственно развитие неустойчивости.

В нашем случае, т.е. при  $\alpha = 1$ , несмотря на то, что компонента скорости  $v_z$  положена равной нулю, компоненты скорости  $v_r$  и  $(v_\varphi/r - \alpha v_z)$  близки к собственным. Особенно это относится к компоненте  $v_r$ , которая в начале развития неустойчивости является определяющей, так как  $(v_\varphi/r - \alpha v_z)$  стоит при производной  $\partial/\partial\theta$ , а в начале нет зависимости от  $\theta$ .

Были рассмотрены следующие комбинации первых четырех мод:

I:	$\lambda_1 = 0,01,$	$\lambda_2 = 0,$	$\lambda_3 = 0,$	$\lambda_4 = 0,$	$N = 1;$
II:	$\lambda_1 = 0,$	$\lambda_2 = 0,01,$	$\lambda_3 = 0,$	$\lambda_4 = 0,$	$N = 1;$
III:	$\lambda_1 = 0,$	$\lambda_2 = 0,$	$\lambda_3 = 0,1,$	$\lambda_4 = 0,$	$N = 1;$
IV:	$\lambda_1 = 0,05,$	$\lambda_2 = 0,05,$	$\lambda_3 = 0,$	$\lambda_4 = 0,$	$N = 1;$
V:	$\lambda_1 = 0,025,$	$\lambda_2 = 0,05,$	$\lambda_3 = 0,$	$\lambda_4 = 0,$	$N = 1;$
VI:	$\lambda_1 = 0,025,$	$\lambda_2 = -0,05,$	$\lambda_3 = 0,$	$\lambda_4 = 0,$	$N = 1;$
VII:	$\lambda_1 = 0,025,$	$\lambda_2 = 0,05,$	$\lambda_3 = 0,075,$	$\lambda_4 = 0,$	$N = 1;$
VIII:	$\lambda_1 = 0,025,$	$\lambda_2 = 0,05,$	$\lambda_3 = 0,075,$	$\lambda_4 = -0,1,$	$N = 1;$
IX:	$\lambda_1 = 0,05,$	$\lambda_2 = 0,05,$	$\lambda_3 = 0,$	$\lambda_4 = 0,$	$N = 2;$
X:	$\lambda_1 = 0,025,$	$\lambda_2 = 0,05,$	$\lambda_3 = 0,075,$	$\lambda_4 = 0,$	$N = 2;$
XI:	$\lambda_1 = 0,025,$	$\lambda_2 = 0,05,$	$\lambda_3 = 0,075,$	$\lambda_4 = -0,1,$	$N = 2;$
XII:	$\lambda_1 = 0,025,$	$\lambda_2 = 0,05,$	$\lambda_3 = 0,075,$	$\lambda_4 = 0,1,$	$N = 2;$
XIII:	$\lambda_1 = 0,025,$	$\lambda_2 = 0,05,$	$\lambda_3 = -0,075,$	$\lambda_4 = -0,1,$	$N = 2;$
XIV:	$\lambda_1 = 0,025,$	$\lambda_2 = 0,05,$	$\lambda_3 = 0,075,$	$\lambda_4 = -0,1,$	$N = 6.$

Заметим, что если рассматривать каждую моду отдельно, то максимальное значение амплитуды возмущения для  $v_r$ :

- при  $m = 1$  равно  $\lambda_1$  (в точке  $r = 0$ );
- при  $m = 2$  равно  $0,38\lambda_2$  (в точке  $r = 0,58$ );
- при  $m = 3$  равно  $0,25\lambda_3$  (в точке  $r = 0,71$ );
- при  $m = 4$  равно  $0,19\lambda_4$  (в точке  $r = 0,78$ ).

Почти все варианты считались с  $\lambda_1 = 0,025$ ,  $\lambda_2 = 0,05$ ,  $|\lambda_3| = 0,075$ ,  $|\lambda_4| = 0,1$ , т.е. амплитуды соответственно равны  $0,025$ ,  $0,019$ ,  $0,019$ ,  $0,019$ . И только для вариантов IV и IX  $v_{r \max 1} = 0,05$  и  $v_{r \max 2} = 0,019$ .

Результаты расчетов приведены на рис. 8–23. Рис. 8, 9–12 относятся к отдельным модам. На рис. 8 показано изменение во времени кинетической энергии  $W_k = \frac{1}{2} \int \rho v^2 dV$  для  $m = 1, 2, 3$  при  $N = 1$ . График энергии имеет характерный вид кривой с максимумом, причем максимальное значение определяется, с одной

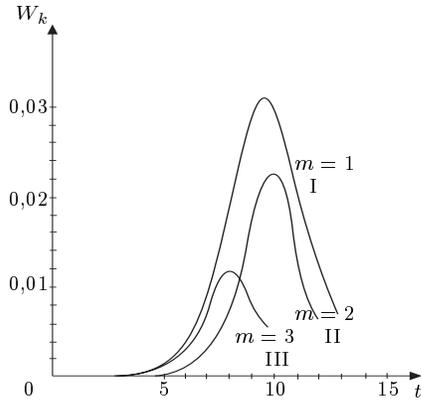


Рис. 8

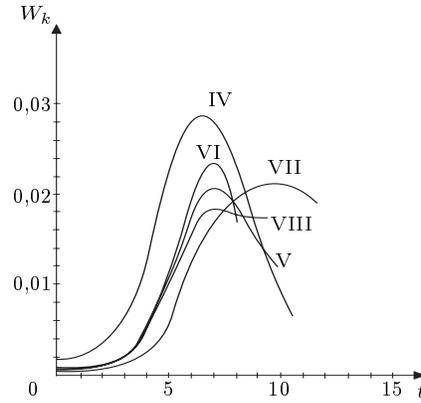


Рис. 9

стороны, параметрами равновесной конфигурации (так, для  $N = 2$  оно в 10 раз больше), а с другой — номером моды (падает с увеличением  $m$ ), и не зависит при малых  $\lambda$  от амплитуды возмущения.

Из рис. 8 можно сделать вывод, что если амплитуды возмущений брать равными, то максимумы для разных мод в нашем случае достигаются приблизительно в одинаковые моменты времени. Это следует из того, что первая мода считалась с  $\lambda_1 = 0,01$ , вторая с  $\lambda_2 = 0,01$ , третья — с  $\lambda_3 = 0,1$ , т.е. максимальное значение  $v_r$  — компоненты скорости в начальный момент в первом случае равнялось  $0,01$ , во втором  $0,0038$ , в третьем  $0,025$ . А как показывают расчеты, уменьшение или увеличение возмущения ведет к более медленному или более быстрому развитию неустойчивости.

На рис. 10–12 показана эволюция во времени нормальных сечений магнитных поверхностей  $\psi(r, \theta) = \text{const}$ , которые при  $t = 0$  были окружностями (для  $m = 1, 2, 3$ ). Видно, что в результате развития неустойчивости магнитные поверхности внутри некоторого круга (“круга неустойчивости”) претерпевают существенное изменение, хотя и не меняют топологии. Плазма как бы выворачивается наизнанку, так что внутренние слои выходят на периферию “круга неустойчивости”, а внешние, наоборот, попадают внутрь. В результате образуются

характерные нелинейные моды с узкими инверсными слоями, характеризующиеся большой плотностью электрического тока.

Это справедливо для всех рассматриваемых мод, причем радиус “круга неустойчивости” для этих мод один и тот же  $\approx 0,7$ . Последний результат согласуется с линейной теорией. На рис. 8, а приведены радиальные собственные функции  $v_r$  для рассматриваемых вариантов. Область локализации приблизительно 0,7 для всех мод. Как показывают наши расчеты, именно область локализации собственной функции и определяет “круг неустойчивости”. Вероятно, на очень больших временах в неустойчивость могут вовлекаться и более дальние слои плазмы.

Отметим, что неустойчивость моды  $m = 1$  отличается от  $m \geq 2$  смещением магнитной оси к границе “круга неустойчивости”.

На рис. 9, 13–17 представлены результаты, относящиеся к параболическому току, т.е.  $N = 1$ , в случае взаимодействия нескольких мод. На рис. 9 показано изменение кинетической энергии для вариантов IV–VIII. Для графика энергии характерно такое же поведение, как и в случае отдельной моды, причем максимумы кривых энергии заключены между максимумами энергии отдельных мод.

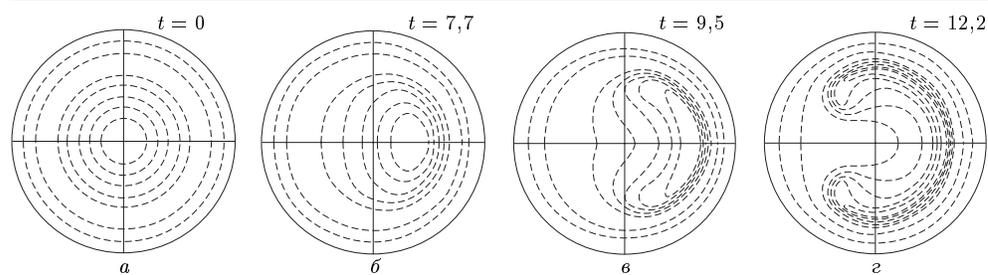
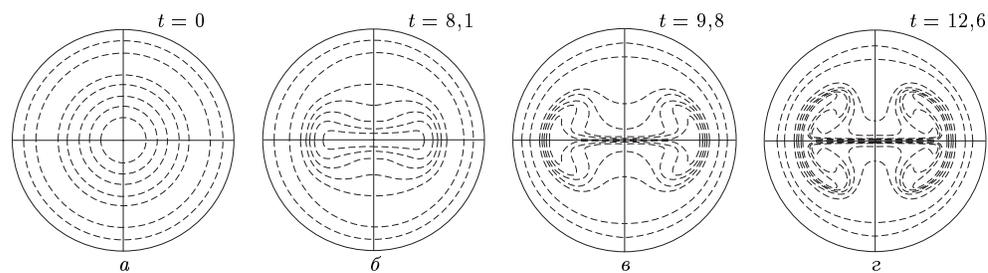
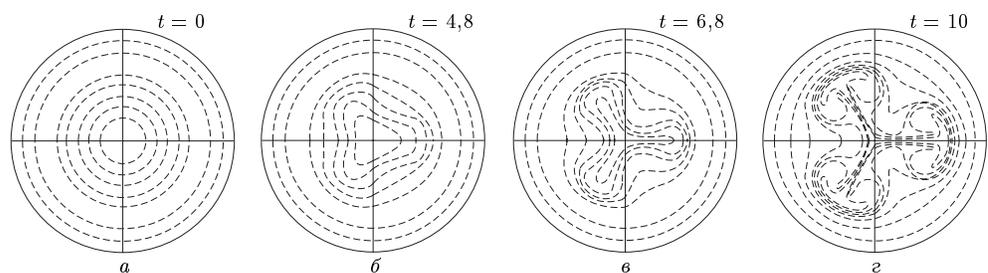
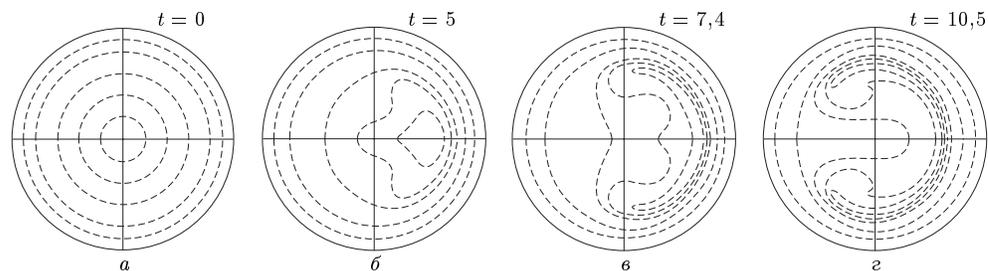
Эволюция сечений магнитных поверхностей для этих вариантов отражена на рис. 13–17. Характерно, что на временах  $\approx 5$  в развитии неустойчивости проявляется влияние всех рассматриваемых мод. Причем очертания магнитных поверхностей (за исключением варианта VI) носят черты либо третьей моды (варианты IV, V), либо третьей и четвертой моды (варианты VII, VIII).

На больших временах в варианте IV (в котором амплитуда возмущения для первой моды в 2,5 раза больше, чем второй) и в варианте VII (в котором действие первой моды усиливалось третьей) первая мода забивает все остальные.

А в варианте V и в варианте VI налицо взаимодействие обеих мод и на больших временах. Так в варианте V есть сдвиг магнитной оси, характерный для первой моды, и образование узкого слоя на оси  $(0, \pi)$  с двумя провалами, характерного для второй.

На рис. 18–22 показана эволюция сечений магнитных поверхностей для вариантов IX–XIII. Эти варианты считались для  $N = 2$ . Согласно линейной теории в этом случае инкременты винтовой неустойчивости больше, чем для  $N = 1$ , причем существенно, что большим модам соответствуют большие инкременты и, кроме того, в развитие неустойчивости захватывается большая область.

Результаты расчета нелинейной задачи целиком согласуются с этим выводом. Неустойчивость захватывает почти всю область до кожуха, причем ее развитие происходит вдвое быстрее, чем для  $N = 1$ . В суммарном движении преобладающую роль играют третья и четвертая мода. Это проявляется не только на временах  $\approx 2,5$ , но и, в отличие от случая  $N = 1$ , на больших временах. Тем не менее существенное влияние первой и второй моды проявляется в заметном сдвиге магнитной оси (первая мода) и симметричном глубоком затекании периферийных слоев к центру (характерно для второй моды). Взаимодействие только первой и второй моды исследовано в варианте IX. Описанные различия в развитии неустойчивости между плазменными конфигурациями с  $N = 1$  и  $N = 2$

Рис. 10. I:  $\psi = \text{const}$ ,  $\alpha = 1$ ,  $\mu_0 = 1$ ,  $N = 1$ ,  $\lambda_1 = 0,001$ Рис. 11. II:  $\psi = \text{const}$ ,  $\alpha = 1$ ,  $\mu_0 = 1$ ,  $N = 1$ ,  $\lambda_2 = 0,001$ Рис. 12. III:  $\psi = \text{const}$ ,  $\alpha = 1$ ,  $\mu_0 = 1$ ,  $N = 1$ ,  $\lambda_3 = 0,1$ Рис. 13. IV:  $\psi = \text{const}$ ,  $\alpha = 1$ ,  $\mu_0 = 1$ ,  $N = 1$ ,  $\lambda_1 = 0,05$ ,  $\lambda_2 = 0,05$

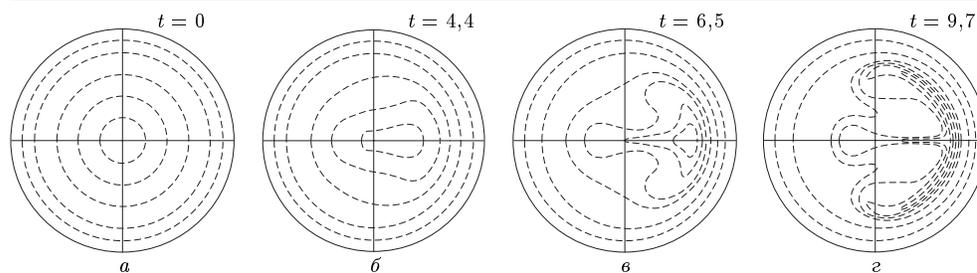


Рис. 14. V:  $\psi = \text{const}$ ,  $\alpha = 1$ ,  $\mu_0 = 1$ ,  $N = 1$ ,  $\lambda_1 = 0,025$ ,  $\lambda_2 = 0,05$

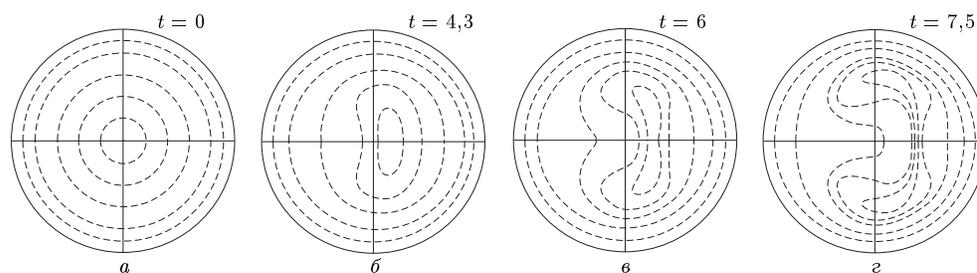


Рис. 15. VI:  $\psi = \text{const}$ ,  $\alpha = 1$ ,  $\mu_0 = 1$ ,  $N = 1$ ,  $\lambda_1 = 0,025$ ,  $\lambda_2 = -0,05$

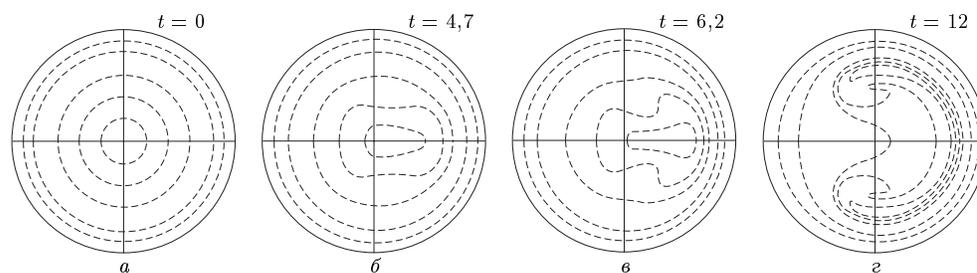


Рис. 16. VII:  $\psi = \text{const}$ ,  $\alpha = 1$ ,  $\mu_0 = 1$ ,  $N = 1$ ,  $\lambda_1 = 0,025$ ,  $\lambda_2 = 0,05$ ,  $\lambda_3 = 0,075$

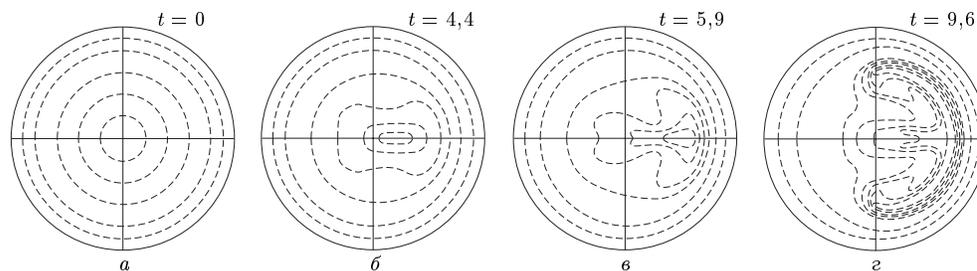


Рис. 17. VIII:  $\psi = \text{const}$ ,  $\alpha = 1$ ,  $\mu_0 = 1$ ,  $N = 1$ ,  $\lambda_1 = 0,025$ ,  $\lambda_2 = 0,05$ ,  $\lambda_3 = 0,075$ ,  $\lambda_4 = -0,01$

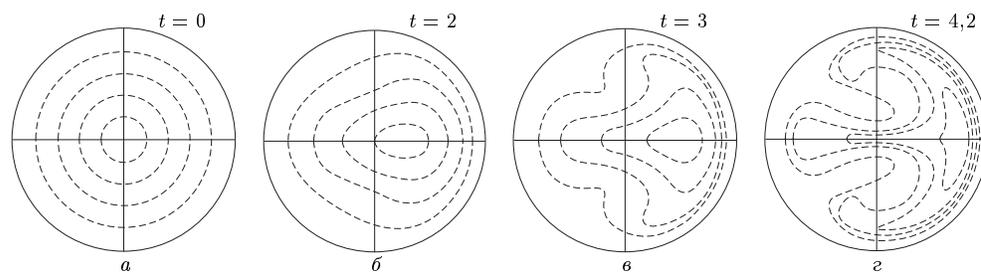


Рис. 18. IX:  $\psi = \text{const}$ ,  $\alpha = 1$ ,  $\mu_0 = 1$ ,  $N = 1$ ,  $\lambda_1 = 0,025$ ,  $\lambda_2 = 0,05$ ,  $\lambda_3 = 0,075$ ,  $\lambda_4 = -0,01$

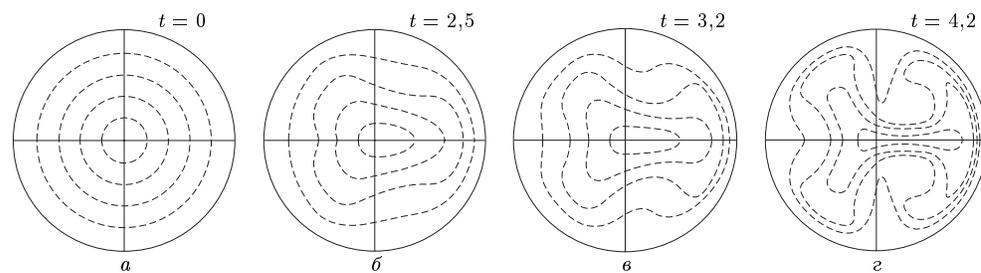


Рис. 19. X:  $\psi = \text{const}$ ,  $\alpha = 1$ ,  $\mu_0 = 1$ ,  $N = 2$ ,  $\lambda_1 = 0,025$ ,  $\lambda_2 = 0,05$ ,  $\lambda_3 = 0,075$

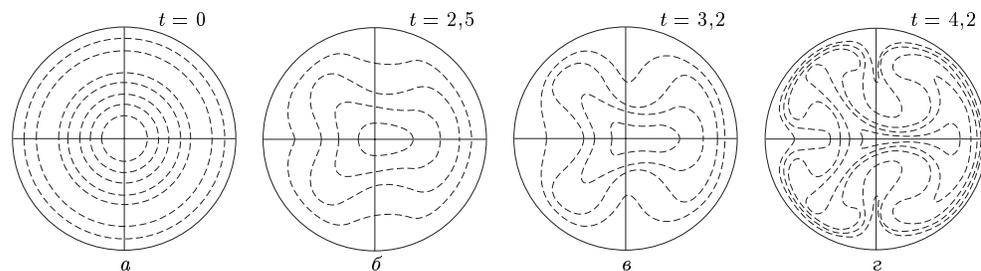


Рис. 20. XI:  $\psi = \text{const}$ ,  $\alpha = 1$ ,  $\mu_0 = 1$ ,  $N = 2$ ,  $\lambda_1 = 0,025$ ,  $\lambda_2 = 0,05$ ,  $\lambda_3 = 0,075$ ,  $\lambda_4 = -0,1$

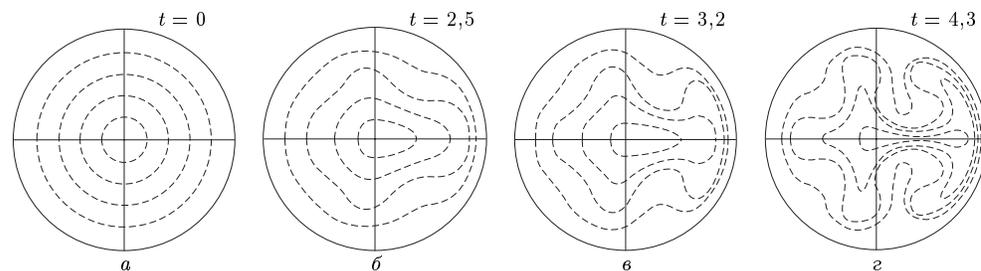


Рис. 21. XII:  $\psi = \text{const}$ ,  $\alpha = 1$ ,  $\mu_0 = 1$ ,  $N = 2$ ,  $\lambda_1 = 0,025$ ,  $\lambda_2 = 0,05$ ,  $\lambda_3 = 0,075$ ,  $\lambda_4 = 0,1$

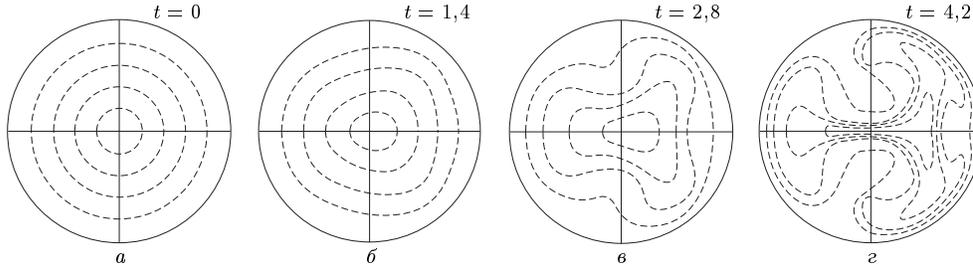


Рис. 22. XIII:  $\psi = \text{const}$ ,  $\alpha = 1$ ,  $\mu_0 = 1$ ,  $N = 2$ ,  $\lambda_1 = 0,025$ ,  $\lambda_2 = 0,05$ ,  $\lambda_3 = -0,075$ ,  $\lambda_4 = -0,1$

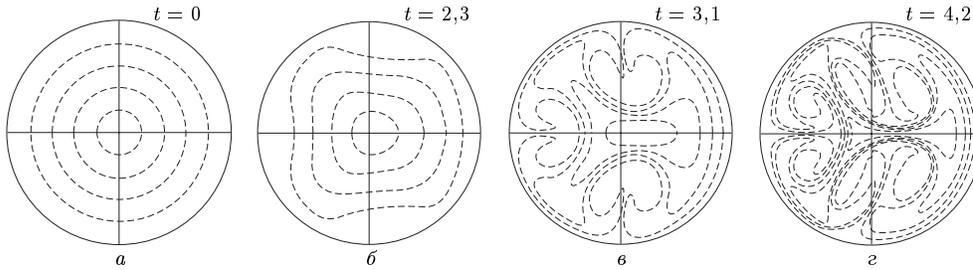


Рис. 23. XIV:  $\psi = \text{const}$ ,  $\alpha = 1$ ,  $\mu_0 = 1$ ,  $N = 6$ ,  $\lambda_1 = 0,025$ ,  $\lambda_2 = 0,05$ ,  $\lambda_3 = 0,075$ ,  $\lambda_4 = -0,1$

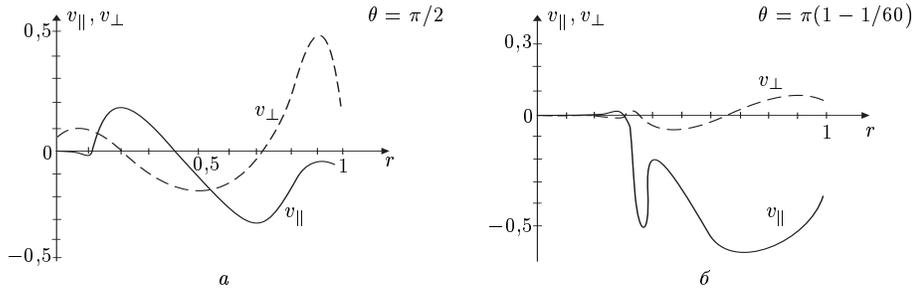


Рис. 24

отчетливо видны при сравнении варианта IV с вариантом IX и варианта VII с вариантом X (см. рис. 13, рис. 18, рис. 16, рис. 19).

Еще сильнее развивается неустойчивость при  $N = 6$ , когда ток становится почти однородным. Это видно при сравнении варианта XI с вариантом XIV (см. рис. 20 и рис. 23). Здесь в движение вовлечена вся плазма вплоть до кожуха.

Таким образом, при однородном токе становится существенным влияние кожуха, при параболическом же токе с резонансом на магнитной оси влияние границы на развитие неустойчивости несущественно, так как возмущенное движение практически не выходит за круг неустойчивости.

Когда в ходе деформации поверхностей возникают узкие слои с большими градиентами магнитного поля, частицы плазмы заметно ускоряются. На рис. 24 приведены графики скоростей

$$v_{\parallel} = \frac{v_z + \alpha r v_{\varphi}}{\sqrt{1 + \alpha^2 r^2}}, \quad v_{\perp} = \frac{v_{\varphi} - \alpha r v_z}{\sqrt{1 + \alpha^2 r^2}}$$

в конфигурации рис. 23,  $z$  вдоль лучей  $\theta = \pi/2$  и  $\theta = \pi(1 - 1/60)$ . Видно, что в местах, где луч пересекает узкий слой, скорость  $v_{\parallel}$  вдоль винтовой линии резко увеличена, а скорость поперек близка к нулю. Это показывает, что узкие слои сильно ускоряют частицы плазмы по касательной к магнитным поверхностям.

### СПИСОК ЛИТЕРАТУРЫ

1. Кадомцев Б.Б. Гидромагнитная устойчивость плазмы // Вопросы теории плазмы. Вып. 2. — М., Госатомиздат, 1963. — С. 132–176.
2. Шафранов В.Д. К вопросу о гидромагнитной устойчивости плазменного шнура с током в сильном продольном магнитном поле // ЖТФ. 1970. Т. 40, вып. 2. С. 241–254.
3. Соловьев Л.С. Гидромагнитная устойчивость замкнутых плазменных конфигураций // Вопросы теории плазмы. Вып. 6. — М.: Атомиздат, 1972. — С. 210–290.
4. Кадомцев Б.Б., Погуце О.П. Нелинейные винтовые возмущения плазмы в токамаке // ВЭТФ. 1973. Т. 65, вып. 2. С. 575–589.
5. Bateman G., Schneider W., Grossman W. MHD-instabilities as an initial boundary-value problem // Nuclear Fusion. 1974. V. 14, № 5. P. 669–683.
6. Wotten J., Hichs H.R., Bateman G., Dory R.A. // Oak Ridge National Laboratory Report, ORNL-TM-4784, 1974.
7. Герлах Н.И., Зуева Н.М. Результаты расчета динамики винтовой МГД-неустойчивости. — М., 1974. — (Препр. / ИПМ АН СССР; № 91).
8. Зуева Н.М., Соловьев Л.С. О винтовой неустойчивости плазменного шнура. — М., 1974. — (Препр. / ИПМ АН СССР; № 89).
9. Bateman G., Hichs H.R., Wooten J.W., Dory R.A. // Proc. 7th European Conf. on Controlled Fusion and Plasma Physics. 1975. V. 1. P. 110.
10. Герлах Н.И., Зуева Н.М., Морозов А.И., Соловьев Л.С. О развитии винтовой МГД-неустойчивости // Письма в ЖТФ. 1975. Т. I, вып. 22. С. 1016–1019; М., 1975. — (Препр. / ИПМ АН СССР; № 74).
11. Зуева Н.М., Соловьев Л.С. Изменение геометрии магнитного поля в движущейся плазме // Письма в ЖТФ. 1975. Т. 1, вып. 24. С. 1117–1122; М., 1975. — (Препр. / ИПМ АН СССР; № 95).
12. Герлах Н.И., Зуева Н.М., Соловьев Л.С. Винтовая МГД-неустойчивость идеальной плазмы. — М., 1975. — (Препр. / ИПМ АН СССР; № 96).
13. Герлах Н.И., Зуева Н.М., Соловьев Л.С. Винтовая неустойчивость при конечной проводимости. — М., 1975. — (Препр. / ИПМ АН СССР; № 111).
14. Данилов А.Ф., Днестровский Ю.Н., Костомаров Д.П., Попов А.М. Нелинейные винтовые волны в плазме с учетом конечной проводимости // Физика плазмы. 1976. Т. 2, вып. 1. С. 167–170.
15. Wesson J.A. Non-linear behaviour of hydromagnetic instabilities // Computer Physics Communications. 1976. V. 12. P. 53–65.
16. Waddel B.V., Rosenbluth M.N., Monticello D.A., White R.B. Non-linear growth of the  $m = 1$  tearing mode // Nuclear Fusion. 1976. V. 16, № 3. P. 528–532.
17. Herrnegger F., Schneider W. Ideal MHD stability of  $m = 2$  modes in diffuse high- $\beta$ ,  $l = 1$  equilibria // Nuclear Fusion. 1976. V. 16, № 6. P. 925–936.
18. Sykes A., Wesson J.A. // Phys. Rev. Letters. 1976. V. 37. P. 140–143.
19. Itoie S., Iton K., Yoshikawa S. Helical Resonance in Tokamak equilibrium // Journal of the Physical Society of Japan. 1976. V. 40, № 6. P. 1762–1786.
20. Герлах Н.И., Зуева Н.М., Соловьев Л.С. Нелинейная винтовая МГД-неустойчивость. — М., 1976. — (Препр. / ИПМ АН СССР; № 84).
21. Данилов А.Ф., Днестровский Ю.Н., Костомаров Д.П., Попов А.М. Образование и эволюция магнитных островов в плазме токамака // Физика плазмы. 1977. Т. 3, вып. 2. С. 215–219.

22. Зуева Н.М., Соловьев Л.С. Нелинейная винтовая МГД-неустойчивость // Магнитная гидродинамика. Т. 3. 1977. С. 5–12.
23. Герлаз Н.И., Зуева Н.М., Соловьев Л.С. Линейная теория винтовой МГД-неустойчивости. — М., 1977. — (Препр. / ИПМ АН СССР; № 62).
24. Зуева Н.М., Палейчик В.В., Соловьев Л.С. Нелинейная винтовая МГД-неустойчивость моды  $m = 1$  для идеальной проводимости. — М., 1977. — (Препр. / ИПМ АН СССР; № 118).
25. Днестровский Ю.Н., Костомаров Д.П., Переверзев В.Т., Тарасян К.Н. Перезамыкание магнитных поверхностей и перемешивание тока в начальной стадии разряда в токамаке // Физика плазмы. 1978. Т. 4, вып. 5. 1001–1003.
26. Соловьев Л.С. Симметричные МГД-течения и винтовые волны в круглом плазменном цилиндре // Вопросы теории плазмы. Вып. 3. — М., Госатомиздат, 1963. — С. 245–289.
27. Сайдем Б. Устойчивость самосжатого линейного разряда // Труды Второй международной конференции по мирному использованию атомной энергии. Т. 1. — М.: Атомиздат, 1959. — С. 89–93.
28. Мирнов С.В., Семенов И.Б. Наблюдение тонкой структуры неустойчивости срыва в токамаке. — М., 1976. — (Препр. / ИАЭ; № 2723).
29. Зуева Н.М., Михайлова М.С., Соловьев Л.С. О конвективной неустойчивости газового шара // Письма в ЖЭТФ. 1977. Т. 26, вып. 3. С. 165–168; М., 1977. — (Препр. / ИПМ АН СССР; № 65).
30. Зуева Н.М., Палейчик В.В., Соловьев Л.С. О развитии конвективной неустойчивости во вращающемся газе // Письма в ЖТФ. 1978. Т. 4, вып. 8. С. 436–440; М., 1978. — (Препр. / ИПМ АН СССР; № 11).
31. Келдыш М.В. О полноте собственных функций некоторых классов несамосопряженных линейных операторов // Успехи матем. наук. 1971. Т. 26, вып. 4. С. 15–41; Доклады АН СССР. 1951. Т. 77, № 1. С. 11–14.
32. Ватсон Г.Н. Теория бесселевых функций. Т. 1. — М.: ИЛ, 1949.
33. Ландау Л.Д., Лифшиц Е.М. Механика сплошных сред. — 2-е изд. — М.: Гостехиздат, 1954.

## ВЕКТОРНО-СТОХАСТИЧЕСКИЕ МОДЕЛИ НЕКОТОРЫХ ЗАДАЧ МАТЕМАТИЧЕСКОЙ ФИЗИКИ \*)

Известный алгоритм метода Монте-Карло для решения задачи Дирихле для уравнения Лапласа [1] — метод “блуждания по сферам” — обобщен на ряд уравнений второго порядка (см., например, [2, 3]) и на некоторые уравнения более высокого порядка с постоянными коэффициентами [4]. Все эти обобщения основаны на возможности перехода от исходного дифференциального уравнения к специальному интегральному уравнению 2-го рода с вырожденным ядром. Такой подход применим к широкому классу эллиптических уравнений [5], для которых выполняется принцип выметания [6]. Хотя, в отличие от уравнений второго

\*) Актуальные проблемы прикладной математики и математического моделирования. — Новосибирск: Наука, 1982. — С. 69–82 (совм. с Г.А. Михайловым, К.К. Сабельфельдом).

порядка, этот оператор для уравнений высокого порядка может иметь отрицательную часть, метод Монте-Карло применим и в этом случае (такой пример подробно рассматривается в п. 4).

В настоящей работе обосновывается алгоритм метода Монте-Карло для решения метагармонических уравнений, исследуется дисперсия этого алгоритма и обсуждаются некоторые возможности применения предлагаемого метода к решению задач теории пластин и некоторых задач на собственные значения.

**1. Постановка задачи и основные определения.** В области  $G \subset R^m$  с границей  $\partial G$  рассмотрим уравнение

$$R_\gamma(\Delta)u(x) = 0, \quad x \in G, \quad (1.1)$$

с граничными условиями

$$\begin{aligned} \Delta^k u(x) &\rightarrow \varphi_k(x_0), \quad x \rightarrow x_0 \in \partial G, \quad k = 0, 1, \dots, \gamma - 1, \\ \Delta^0 u &\equiv u. \end{aligned} \quad (1.2)$$

Здесь  $R_\gamma(\lambda) = \lambda^\gamma + c_1 \lambda^{\gamma-1} + \dots + c_{\gamma-1} \lambda + c_\gamma$  — многочлен с вещественными коэффициентами;  $\Delta$  — оператор Лапласа.

Обозначим через  $\{\lambda_k\}_{k=1}^\gamma$  корни уравнения  $R_\gamma(\lambda) = 0$ . Будем предполагать, что выполнены условия однозначной разрешимости задачи (1.1), (1.2). Для этого потребуем, чтобы граница  $\partial G$  была регулярной [4],  $\{\varphi_k\}_{k=1}^\gamma$  — непрерывные функции на границе области  $G$  и  $\operatorname{Re} \lambda_i \neq \mu_k$ ,  $i = 1, \dots, \gamma$ ,  $k = 1, 2, \dots$ , где  $\{\mu_k\}_{k=1}^\infty$  — собственные значения оператора Лапласа. Последнее условие имеет место, например, если  $\operatorname{Re} \lambda_i \geq 0$ ,  $i = 1, \dots, \gamma$ . В дальнейшем при рассмотрении алгоритма метода Монте-Карло, использующего  $\varepsilon$ -смещение граничных значений, будут сделаны дополнительные предположения о гладкости решения.

Рассмотрим однородное уравнение (1.1). Обобщение на неоднородный случай с правой частью  $f(x)$ ,  $x \in G$ , не представляет принципиальных трудностей и обсуждается ниже. Предположим, что необходимо найти решение задачи (1.1), (1.2) в точке  $x_0 \in G$ . Пусть  $W(t)$  — винеровский процесс в  $R^m$  (сужением инфинитезимального оператора которого является оператор Лапласа), выходящий из точки  $x_0$ :  $W(0) = x_0$ . Обозначим через  $\tau$  момент первого выхода процесса  $W(t)$  из области  $G$ . Для задачи (1.1), (1.2) (с правой частью  $f(x)$ ) существует вероятностное представление решения в виде винеровского интеграла [7]:

$$u(x_0) = \sum_{i=1}^{\gamma-1} \mathbf{M}_{x_0}[y_i(\tau) \varphi_i(X_\tau)] + \mathbf{M}_{x_0} \int_0^\tau y_{\gamma-1}(t) f(X_t) dt, \quad (1.3)$$

где  $W(\tau) = X_\tau$ , а  $y_i(s)$  — решение уравнения

$$y^{(\gamma)} - c_1 y^{(\gamma-1)} + \dots + (-1)^\gamma c_\gamma y = 0,$$

удовлетворяющее условиям:  $y_i^{(j)} = 0$  при  $i \neq j$  и  $y_i^{(i)} = 1$ ,  $i, j = 0, 1, \dots, \gamma - 1$ . Представление (1.3) верно, если уравнение  $\Delta u + \mu u = 0$  имеет в области  $\bar{G}$  положительное решение, где  $\mu = \max_i \operatorname{Re} \mu_i$ ;  $\mu_i$  — корни уравнения  $R_\gamma(-\mu) = 0$ .

Для вычисления соответствующих винеровских интегралов в (1.3) можно, таким образом, использовать метод Монте-Карло [8–10], при этом  $\tau$  моделируется по известному алгоритму [9].

В случае если все корни уравнения  $R_\gamma(\lambda) = 0$  различны, задачу можно расщепить на  $\gamma$  задач Дирихле для уравнения Пуассона [10]. Действительно, решение задачи (1.1), (1.2) можно тогда искать в виде

$$u(x) = \sum_{j=0}^{\gamma-1} u_j, \quad \Delta u_j - \lambda_j u_j = 0, \quad (1.4)$$

$$u_j|_{\partial G} = \sum_{i=0}^{\gamma-1} c_{ij} \varphi_i(X_\tau), \quad (1.5)$$

где коэффициенты  $c_{ij}$  выражаются через  $c_i$ ,  $i = 1, \dots, \gamma$ . Решение задач (1.4), (1.5) может быть одновременно для всех  $j = 1, \dots, \gamma$  получено методом “блуждания по сферам” [2]. Поскольку, однако, для приложений наиболее интересным является случай бигармонического уравнения, а также общего полигармонического уравнения, изложим здесь другой подход к решению этой задачи, основанный на использовании одного обобщения известной теоремы о среднем значении для гармонических функций.

**2. Построение векторного алгоритма метода Монте-Карло.** Рассмотрим некоторую ограниченную область  $G \subset R^m$  с границей  $\partial G$ . Для точек  $x \in G$  определим функцию  $d(x) = \inf_{y \in \partial G} |x - y|$  — расстояние от точки  $x$  до границы  $\partial G$ . Определим также  $\varepsilon$ -окрестность границы  $\partial G$ :

$$\partial G_\varepsilon = \{x \in \bar{G} : d(x) < \varepsilon\}.$$

Пусть  $S(x_0, r)$  есть сфера радиуса  $r$  с центром в точке  $x_0$  и пусть  $S(x_0, r) \subset \bar{G}$ .

Введем оператор усреднения

$$N(u) \equiv N(u(r)) = \int_{S(x_0, r)} u(x) d\sigma_{x_0}^r(x), \quad (2.1)$$

где  $\sigma_{x_0}^r$  — мера Лебега на  $S(x_0, r)$ . Алгоритм метода Монте-Карло для решения задачи (1.1), (1.2) основан на следующем утверждении [4].

**ТЕОРЕМА 1.** *Любое решение уравнения (1.1) удовлетворяет системе интегральных уравнений*

$$\Delta^{l-1} u(x_0) = \sum_{j=1}^{\gamma} c_{lj}(r) \int_{S(x_0, r)} \Delta^{j-1} u(x) d\sigma_{x_0}^r(x), \quad l = 1, 2, \dots, \gamma. \quad (2.2)$$

Коэффициенты  $c_{lj}(r)$  явно выражаются через функции Бесселя  $\{I_{m/2-1}(r\sqrt{\lambda_j})\}_{j=1, \dots, \gamma}$ .

Так как доказательство этой теоремы подробно изложено в [4], мы приводим здесь лишь основную схему доказательства. Рассмотрим вспомогательное уравнение

$$L(u) - \lambda u = 0, \quad (2.3)$$

где  $L$  — произвольный линейный оператор,  $\lambda$  — константа. Предположим, что операторы  $N(u)$  и  $L(u)$  перестановочны, тогда  $L(N(u)) - \lambda N(u) = 0$ , т.е.  $N(u)$  также является решением уравнения (2.3). Заметим, что  $N(u)$  зависит только от  $r$  (при фиксированном  $x_0$ ), поэтому, переходя к полярной системе координат, получим обыкновенное дифференциальное уравнение, решение которого дает нам представление для  $N(u)$ . В случае когда  $L = \Delta$  (оператор Лапласа перестановочен с  $N$ ), уравнение (2.3) переходит в

$$\frac{d^2 u}{dr^2} + \frac{m-1}{r} \frac{du}{dr} - \lambda u = 0, \quad (2.4)$$

поэтому

$$N(u) = C W_m(r, \lambda) + D V_m(r, \lambda), \quad (2.5)$$

где  $W_m(r, \lambda), V_m(r, \lambda)$  — линейно независимые решения уравнения (2.4). Из (2.5) получается известное соотношение о среднем для уравнения Гельмгольца. Действительно, так как  $N(u) \in C^2$  в точке  $r = 0$  (см., например, [4]), имеем  $D = 0$ . Следовательно,  $C = N(u)|_{r=0} = u(x_0)$ , поэтому  $N(u) = u(x_0) W(r, \lambda)$ , где

$$\begin{aligned} W(r, \lambda) \equiv W_m(r, \lambda) &= \Gamma\left(\frac{m}{2}\right) 2^{m/2-1} \frac{I_{m/2-1}(r\sqrt{\lambda})}{(r\sqrt{\lambda})^{m/2-1}} = \\ &= \sum_{k=0}^{\infty} \frac{(r^2 \lambda)^k}{2 \cdot 4 \cdot \dots \cdot 2k \cdot m(m+2) \cdot \dots \cdot (m+2k-2)}. \end{aligned} \quad (2.6)$$

Функция  $W(r, \lambda)$  и ее производные  $\{\partial^k W / \partial \lambda^k\}_{k=1,2,\dots}$  являются целыми функциями. Через эти функции выражаются коэффициенты  $c_{ij}(r)$  в (2.2), причем от производных можно избавиться, так как

$$\begin{aligned} \frac{\partial^k W_m(r, \lambda)}{\partial \lambda^k} &= \Gamma\left(\frac{m}{2}\right) \lambda^{-k} \left(\frac{r\sqrt{\lambda}}{2}\right)^{k-m/2+1} I_{m/2-k-1}(r\sqrt{\lambda}) = \\ &= \frac{r^{2k}}{2^k m(m+2) \dots (m+2k-2)} W_{m+2k}(r, \lambda). \end{aligned} \quad (2.7)$$

Для произвольного решения уравнения (1.1) имеет место представление [4]

$$N(u) = \sum_{j=1}^k \sum_{i=1}^{m_j} \left\{ \frac{F_i}{(i-1)!} \frac{\partial^{i-1} W}{\partial \lambda^{i-1}} \Big|_{\lambda=\lambda_j} + \frac{D_i}{(i-1)!} \frac{\partial^{i-1} V}{\partial \lambda^{i-1}} \Big|_{\lambda=\lambda_j} \right\}, \quad (2.8)$$

где  $\{\lambda_j\}_{j=1}^k$  — корни (кратности  $m_j$ ,  $m_1 + \dots + m_k = \gamma$ ,  $\lambda_i \neq \lambda_j$  для  $i \neq j$ ) уравнения  $R_\gamma(\lambda) = 0$ ;  $F_i, D_i$  — произвольные константы. Положим

$$B_j(u) = (\Delta - \lambda_1)^{m_1} \dots (\widehat{\Delta - \lambda_j})^{m_j} \dots (\Delta - \lambda_k)^{m_k},$$

где крышечка над сомножителем  $(\Delta - \lambda_j)^{m_j}$  означает, что он отсутствует в этом произведении. Имеет место представление [4]

$$N(u) = \sum_{j=1}^k \sum_{i=1}^{m_j} \left\{ \frac{B_j(\Delta - \lambda_j)^{i-1}}{(\lambda_j - \lambda_1)^{m_1} \dots (\lambda_j - \lambda_k)^{m_k}} \frac{\partial^{i-1} W(r, \lambda)}{\partial \lambda^{i-1}} \Big|_{\lambda=\lambda_j} u \Big|_{r=0} \right\}. \quad (2.9)$$

Из соотношения (2.9) уже нетрудно получить систему интегральных уравнений вида (2.2). Действительно, записывая уравнение (2.9) для  $N(\Delta u)$ ,  $N(\Delta^2 u), \dots, N(\Delta^{\gamma-1} u)$ , получим линейную систему из  $\gamma$  уравнений, откуда придем к системе (2.2). Выпишем здесь систему вида (2.2) для случая, когда  $m_1 = m_2 = \dots = m_k = 1$ . Систему (2.9) удобнее записать в виде

$$\begin{vmatrix} N(u) & W(r, \lambda_1) & \dots & W(r, \lambda_\gamma) \\ u(x_0) & 1 & \dots & 1 \\ \Delta u(x_0) & \lambda_1 & \dots & \lambda_\gamma \\ \dots & \dots & \dots & \dots \\ \Delta^{\gamma-1} u(x_0) & \lambda_1^{\gamma-1} & \dots & \lambda_\gamma^{\gamma-1} \end{vmatrix} = 0.$$

Используя аналогичные уравнения для  $\{N(\Delta^k u)\}_{k=1}^{\gamma-1}$ , получим окончательно

$$\Delta^{l-1} u(x_0) = \sum_{j=1}^{\gamma} c_{lj}(r) N(\Delta^{j-1} u), \quad l = 1, 2, \dots, \gamma, \quad (2.10)$$

где

$$c_{lj}(r) = \frac{1}{\Delta_\gamma} \sum_{i=1}^j \Delta_{ij} \lambda_i^{l-1},$$

$$\Delta_{ij} = \prod_{k \neq i} W(r, \lambda_k) \Delta_{ji}^{(\gamma-1)}, \quad (2.11)$$

$$\Delta_{ji}^{(\gamma-1)} = (-1)^{i+j+1} \prod_{\substack{m \geq p > k \geq 1 \\ k, p \neq i}} (\lambda_p - \lambda_k) \sigma_{m-j}^{(i)}.$$

Здесь  $\sigma_s^{(i)}$  — сумма всевозможных произведений чисел  $\lambda_1, \lambda_2, \dots, \lambda_{i-1}, \lambda_{i+1}, \dots, \lambda_s$ ;

$$\Delta_\gamma = \prod_{\gamma \geq i > k \geq 1} (\lambda_i - \lambda_k) \prod_{j=1}^i W(r, \lambda_j).$$

Выпишем выражения для коэффициентов  $c_{lj}(r)$  для важного в теории пластин случая  $\lambda_1 = \lambda_2 = \lambda$  в общем случае  $m$ -мерного пространства:

$$c_{11} = \frac{W_m(r, \lambda) + \lambda r^2 W_{m+2}(r, \lambda)/2m}{W_m^2(r, \lambda)},$$

$$c_{12} = -r^2 \frac{W_{m+2}(r, \lambda)}{2m} W_m^2(r, \lambda),$$

$$c_{21} = r^2 \lambda^2 \frac{W_{m+2}(r, \lambda)}{2m} W_m^2(r, \lambda),$$

$$c_{22} = \frac{W_m(r, \lambda) - \lambda r^2 W_{m+2}(r, \lambda)/2m}{W_m^2(r, \lambda)}. \quad (2.12)$$

Приступим теперь к построению алгоритма для решения задачи (1.1), (1.2). Предположим, что нам необходимо получить решение задачи (1.1), (1.2) в точке  $x_0 \in G$ . Для функций  $v_i = \Delta^{i-1} u$ ,  $i = 1, \dots, \gamma$ , можно записать ввиду (2.2)

следующую систему интегральных уравнений:

$$v_i(x_0) = \sum_{j=1}^{\gamma} \int_G k_{ij}(x', x_0) v_j(x') dx' + g_i(x_0), \quad (2.13)$$

где

$$g_i(x) = \begin{cases} 0, & x \in G \setminus \partial G_\varepsilon, \\ \Delta^{i-1} u, & x \in \partial G_\varepsilon, \end{cases} \quad (2.14)$$

$$k_{ij}(x', x) = \begin{cases} c_{ij}(r) \delta_x(x'), & x \in G \setminus \partial G_\varepsilon, \\ 0, & x \in \partial G_\varepsilon. \end{cases}$$

Здесь  $\delta_x(x')$  — обобщенная плотность, соответствующая равномерному распределению вероятностей на сфере  $S(x_0, r)$ . Соотношение (2.13) можно рассматривать как сопряженное интегральное уравнение 2-го рода с обобщенным ядром. В данном случае из точки  $x_0$  следует переходить на сферу  $S(x_0, r)$ , т.е. осуществлять “блуждание по сферам”.

Векторную случайную оценку  $\{z_l^{(\varepsilon)}\}_{l=1}^{\gamma}$  решения системы интегральных уравнений (2.13) можно построить следующим образом. Пусть  $\{\xi_k\}$  — процесс “блуждания по сферам” с началом в точке  $x_0$ :  $\xi_k = \xi_{k-1} + \omega_k d(\xi_{k-1})$ ,  $k = 1, 2, \dots$ , где  $\omega_k$  — последовательность изотропных векторов единичной длины. Процесс обрывается, если точка попадает в  $\partial G_\varepsilon$ . Определим векторную случайную величину:

$$z_l^{(\varepsilon)}(\xi_k) = \sum_{j=1}^{\gamma} c_{lj}(r_{k+1}) z_j^{(\varepsilon)}(\xi_{k+1}). \quad (2.15)$$

Здесь  $r_{k+1} = d(\xi_k)$ ,  $l = 1, \dots, \gamma$ ,  $\{z_l^{(\varepsilon)}(x_N) = \varphi_{l-1}(x_N^*)\}_{l=1}^{\gamma}$ , где  $x_N \in \partial G_\varepsilon$  — точка, где произошел обрыв траектории, а  $x_N^* \in \partial G$  — ближайшая к ней граничная точка:  $|x_N - x_N^*| = d(x_N)$ .

Пусть  $C(r) = \|c_{ij}(r)\|_{i,j=1}^{\gamma}$ . Положим

$$C^{(n)} = \|c_{ij}^{(n)}(r_1, \dots, r_n)\|_{i,j=1}^{\gamma} \equiv C(r_0) \prod_{k=1}^n C(r_{k-1} + r_k \omega_k),$$

$$r_k = d(\xi_{k-1}), \quad \|C^{(n)}\| = \sup_{i, \{\xi_0, \omega_k\}} \sum_{j=1}^{\gamma} |c_{ij}^{(n)}|.$$

ТЕОРЕМА 2. 1. Если

$$\rho = \lim_{n \rightarrow \infty} \|C^{(n)}\|^{1/n} = 1, \quad (2.16)$$

то при  $\varepsilon > 0$  спектральный радиус  $\rho_\varepsilon(K) < 1$  и ряд Неймана для системы (2.13) сходится.

2. Условие (2.16) выполняется, если

$$\operatorname{Re} \lambda_i \geq 0, \quad i = 1, 2, \dots, \gamma. \quad (2.17)$$

Доказательство. Обозначим через  $S(P, \varepsilon)$  площадь поверхности  $S(P, d(P))$ . Пусть  $\omega_n(d)$  — площадь поверхности сферы радиуса  $d(P)$  в  $R^n$ . Нам потребуется следующее условие:

$$\min_P \left( \frac{S(P, \varepsilon)}{\omega_n} \right) \geq \nu = \nu(\varepsilon) > 0. \quad (2.18)$$

Нетрудно проверить, что в (2.18) можно положить  $\nu(\varepsilon) = \omega_{m-1}(\varepsilon)/\omega_m(d(P^*))$ .

Рассмотрим  $n$ -ю итерацию оператора  $K$ . Используя оценку

$$\|K\| \leq \sup_{i,x} \sum_{j=1}^{\gamma} |k_{ij}(x, x')| dx'$$

и соотношение (2.18), получим

$$\begin{aligned} \|K^n\| \leq \sup_{i, \{\xi_0, \omega_k\}} & \left\{ \sum_{j=1}^{\gamma} \int_G \delta_{\xi_0}(\xi_1) d\xi_1 \int_{G \setminus \partial G_\varepsilon} \delta_{\xi_1}(\xi_2) d\xi_2 \dots \right. \\ & \left. \dots \int_{G \setminus \partial G_2} \delta_{\xi_{n-1}}(\xi_n) d\xi_n \cdot |c_{ij}^{(n)}| \right\} \leq [1 - \nu(\varepsilon)]^{n-1} \cdot \gamma \|C^{(n)}\|. \end{aligned}$$

Отсюда, учитывая (2.16), получим

$$\lim_{n \rightarrow \infty} \|K^n\|^{1/n} \leq [1 - \nu(\varepsilon)] \cdot \rho < 1.$$

Займемся теперь проверкой условия (2.16). Рассмотрим сначала полигармоническое уравнение  $\Delta^\gamma u = 0$ . В этом случае  $c_{ik}(r) = \alpha_{ik} r^{2(k-i)}$  [5], где

$$\alpha_{ik} = \begin{cases} \beta_{ik} < 1, & k > i, \\ 1, & k = i, \\ 0, & k < i. \end{cases}$$

Поэтому  $\rho = 1$  и условие (2.16) выполняется.

Остановимся теперь подробнее на случае уравнения (1.1) 4-го порядка. Матрица  $C$  имеет вид ( $\lambda_1 \neq \lambda_2$ )

$$C = \frac{1}{\Delta_2} \begin{pmatrix} \lambda_1 W(r, \lambda_1) - \lambda_2 W(r, \lambda_2) & -[W(r, \lambda_1) - W(r, \lambda_2)] \\ \lambda_1 \lambda_2 [W(r, \lambda_1) - W(r, \lambda_2)] & -[\lambda_2 W(r, \lambda_1) - \lambda_1 W(r, \lambda_2)] \end{pmatrix}.$$

Покажем, что собственные значения этой матрицы есть числа  $W_m^{-1}(r, \lambda_1)$ ,  $W_m^{-1}(r, \lambda_2)$ . Положим

$$T_{f(\lambda)} = \begin{pmatrix} 1 & 0 \\ f(\lambda) & 1 \end{pmatrix}, \quad T'_{f(\lambda)} = \begin{pmatrix} 1 & f(\lambda) \\ 0 & 1 \end{pmatrix}.$$

Тогда

$$T_{-\lambda_1} C T_{\lambda_1} = \begin{pmatrix} W_m^{-1}(r, \lambda_1) & -[W_m^{-1}(r, \lambda_1) - W_m^{-1}(r, \lambda_2)]/\Delta_2 \\ 0 & W_m^{-1}(r, \lambda_2) \end{pmatrix}.$$

Так как  $|W(r, \lambda)| \geq 1$  при  $\operatorname{Re} \lambda \geq 0$ , получим  $\rho = 1$ , если  $\operatorname{Re} \lambda_i \geq 0$ ,  $i = 1, 2$ .

Пусть теперь  $\lambda_1 = \lambda_2 = \lambda$ . Матрица  $C$  теперь определяется из (2.12). В этом случае

$$T_{-\lambda} C T_{\lambda} = \begin{pmatrix} W_m^{-1}(r, \lambda) & -r^2 W_{m+2}(r, \lambda) / 2m W_m^2(r, \lambda) \\ 0 & W_m^{-1}(r, \lambda) \end{pmatrix},$$

откуда  $\rho = 1$ , если  $\operatorname{Re} \lambda \geq 0$ . Заметим, что такие преобразования приводят к новому, отличному от (1.4), (1.5) расщеплению задачи. Действительно, в новом базисе при  $\lambda_1 \neq \lambda_2$

$$T'_{1/(\lambda_1 - \lambda_2)} T_{-\lambda_1} C T_{\lambda_1} T'_{1/(\lambda_1 - \lambda_2)} = \begin{pmatrix} W_m^{-1}(r, \lambda_1) & 0 \\ 0 & W_m^{-1}(r, \lambda_2) \end{pmatrix}.$$

Аналогично, в случае различных  $\{\lambda_i\}_{i=1}^{\gamma}$  матрица (2.11) путем линейных преобразований (с коэффициентами, зависящими от  $\lambda_1, \dots, \lambda_{\gamma}$ ) приводится к диагональному виду  $c_{ii} = W_m^{-1}(r, \lambda_i)$ ,  $i = 1, \dots, j$  [5]. В общем случае уравнения (1.1) матрица  $C$  приводится к клеточно-диагональному виду с клетками  $C_j$  [5],

$$C_j = \begin{pmatrix} W_m^{-1}(r, \lambda_j) & c_j^{(1)} r^2 W_{m+2} / W_m^{\gamma} & \dots & c_j^{(j-1)} r^{2(j-1)} W_{m+2(j-1)} / W_m^{\gamma} \\ 0 & W_m^{-1}(r, \lambda_j) & \dots & c_j^{(j-2)} r^{2(j-2)} W_{m+2j} / W_m^{\gamma} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & W_m^{-1}(r, \lambda_j) \end{pmatrix}, \quad (2.19)$$

откуда, учитывая, что  $|W_m(r, \lambda_j)| \geq 1$  при  $\operatorname{Re} \lambda_j \geq 0$ , получим  $\rho = 1$ .

Отметим, что сходимость ряда Неймана еще не обеспечивает применимости метода Монте-Карло, хотя и гарантирует существование и единственность решения соответствующей системы интегральных уравнений. Необходимо еще доказать несмещенность построенной оценки  $z_l(\xi_k)$ . Как ясно из предыдущего, эту оценку можно представить в виде  $z_1 = \xi_1 + \xi_2 + \dots + \xi_k$ , где  $\xi_i$  — оценка для уравнения  $(\Delta - \lambda_i)^{\gamma_i} u = 0$  с преобразованными граничными условиями, причем соответствующая матрица весов  $C^{(i)}$  имеет верхний треугольный вид (2.19) с диагональными элементами  $1/W(\sqrt{\lambda_i} r)$ . Теорема доказана.

**ТЕОРЕМА 3.** Пусть существуют и ограничены нормальные производные  $|\partial v_k / \partial n|_{\partial G_{\varepsilon}} \leq A_k$ , и пусть  $\operatorname{Re} \lambda_i \geq 0$ ,  $i = 1, \dots, \gamma$ . Тогда

$$|\mathbf{M}_{x_0} z_l^{(\varepsilon)} - v_l(x_0)| = 0(\varepsilon), \quad l = 1, \dots, \gamma, \quad (2.20)$$

при  $\varepsilon \rightarrow 0$ .

**Доказательство.** Для доказательства несмещенности  $\mathbf{M}_{x_0} z_l^{(\varepsilon)}$  усредним оценку  $\xi = \xi_1 + \dots + \xi_k$ :  $\mathbf{M}\xi = \mathbf{M}\xi_1 + \dots + \mathbf{M}\xi_k$ . Поэтому достаточно доказать несмещенность  $\mathbf{M}\xi_i$ , т.е. показать, что величина  $\mathbf{M}\xi_i$  дает решение системы интегральных уравнений, соответствующей оператору  $(\Delta - \lambda_i)^{\gamma_i}$ . Для этого достаточно [11] показать, что  $\rho_{\varepsilon}(K_1^{(i)}) < 1$ , где  $K_1^{(i)}$  — интегро-матричный оператор вида (2.2), соответствующий матрице  $C_1^{(i)}$ , элементы которой равны модулям элементов матрицы  $C^{(i)}$ . Вследствие треугольного вида матрицы  $C^{(i)}$  требуемое неравенство получается точно так же, как оценка  $\rho(K) < 1$  в теореме 2. Таким

образом, несмещенность установлена. Для доказательства (2.20) приближенное решение системы (2.13) выпишем в виде ряда Неймана [12]:

$$v_l^{(\varepsilon)}(x_0) = \mathbf{M}_{x_0} \sum_{k=0}^N \sum_{i=1}^{\gamma} g_i(\xi_k) Q_{k,l}^{(i)}, \quad l = 1, \dots, \gamma,$$

где  $Q_{k,l}^{(i)}$  — соответствующий вектор весов, а  $N$  — случайный номер последнего состояния цепи. Используя аналогичное, но точное представление для  $\mathbf{M}_{x_0} z_l^{(\varepsilon)}$ , получим, учитывая сделанное предположение о гладкости решения в  $\partial G_\varepsilon$ ,

$$|v_l^{(\varepsilon)}(x_0) - \mathbf{M}_{x_0} z_l^{(\varepsilon)}| = \left| \mathbf{M} \sum_{i=1}^{\gamma} Q_{N,l}^{(i)} [g_i(x_N^*) - g_i(\xi_N)] \right| = H \cdot \varepsilon,$$

где  $H$  — некоторая константа, зависящая от величин  $A_k$ . Теорема доказана.

Рассмотрим теперь важный вопрос о поведении дисперсии предлагаемого алгоритма.

**3. Исследование дисперсии векторного алгоритма.** Докажем конечность, а в случае выпуклой границы при  $m = 3$  — равномерную ограниченность по  $\varepsilon$  дисперсии изучаемого векторного алгоритма. Для этого достаточно рассмотреть случай, когда дифференциальный оператор имеет вид  $(\Delta - \lambda)^\gamma$ . Соответствующий матричный весовой множитель  $Q_i$  можно представить следующим образом:

$$Q_i \leq E + \mathcal{L} = E + c_i^{(1)} r_i^2 H + c_i^{(2)} r_i^4 H^2 + \dots + c_i^{(\gamma-1)} r_i^{2(\gamma-1)} H^{\gamma-1}.$$

Здесь матрица  $\mathcal{L}$  верхняя треугольная с элементами

$$l_{i,k} = \begin{cases} c_i^{(k)} r_i^{2k}, & k > i, \quad k = 1, \dots, \gamma - 1, \\ 0, & k \leq i, \end{cases}$$

где  $r_i$  — радиус  $i$ -й случайной сферы, а

$$H = \begin{pmatrix} 0 & 1 & & 0 \\ & 0 & 1 & \\ \dots & \dots & \dots & \dots \\ 0 & & 0 & 1 \\ & & & 0 \end{pmatrix},$$

причем  $H^\gamma = 0$  (см., например, [13, с. 25]). Случайный вес после  $n$ -го перехода в цепи “блужданий по сферам” представим в таком же виде:

$$\prod_{i=1}^n Q_i = E + q_n^{(1)} H + q_n^{(2)} H^2 + \dots + q_n^{(\gamma-1)} H^{\gamma-1}.$$

Нетрудно заметить, что  $|c_i^{(k)}| \leq c < \infty$  (причем можно считать, что  $c \geq 1$ ). Поэтому

$$g^{(k)} \leq c^\gamma \sum_{k_1 + \dots + k_n = m} r_1^{2k_1} r_2^{2k_2} \dots r_n^{2k_n} = c^\gamma \sum_{\substack{k_1 \leq k_2 \leq \dots \leq k_n \\ k_1 + \dots + k_n = m}} \sum_{i_1, i_2, \dots, i_n = 1} r_{i_1}^{2k_1} r_{i_2}^{2k_2} \dots r_{i_n}^{2k_n}.$$

Далее,

$$\sum'_{i_1, i_2, \dots, i_n=1}^n r_{i_1}^{2k_1} r_{i_2}^{2k_2} \dots r_{i_n}^{2k_n} = \sum'_{i_1, i_2, \dots, i_n=1}^n r_{i_1}^{2k_1} r_{i_2}^{2k_2} \dots r_{i_n}^{2k_n} + \sum''_{i_1, i_2, \dots, i_n=1}^n r_{i_1}^{2k_1} r_{i_2}^{2k_2} \dots r_{i_n}^{2k_n}.$$

Символ  $\sum'$  здесь означает, что суммирование производится только по группам значений  $i_1, i_2, \dots, i_n$ , среди которых нет совпадающих. Символом  $\sum''$  обозначена сумма по тем группам значений, среди которых есть хотя бы одно, равное  $n$  при отличном от нуля соответствующем значении  $k_i$ . Простые комбинаторные соображения показывают, что

$$\sum''_{i_1, i_2, \dots, i_n=1}^n r_{i_1}^{2k_1} \dots r_{i_n}^{2k_n} \leq cR^{2(k-1)} n^{s-1} r_n^2,$$

где  $R$  — половина диаметра области, а  $s$  — количество фиксированных значений  $k_j$ , отличных от нуля. Таким образом, вопрос об ограниченности дисперсий изучаемых алгоритмов сводится к вопросу об ограниченности дисперсий случайных величин вида

$$\chi_s^{(N)} = \sum_{n=0}^N n^s r_n^2, \quad r = 0, 1, 2, \dots,$$

который и будет рассмотрен далее.

**ЛЕММА.** *Величина  $\mathbf{M}\chi_s^{(\infty)}$  конечна для плоской границы в трехмерном пространстве.*

**Доказательство.** При фиксированном  $r_0$  случайная величина  $r_n$  здесь распределена [9] с плотностью

$$f_n(r) = \frac{[\ln(2^n r_0/r)]^n}{2^n \cdot n! r_0}, \quad r < 2^n r_0.$$

Отсюда

$$\mathbf{M}\chi_s = \frac{1}{r_0} \sum_{n=0}^{\infty} \int_0^{2^n r_0} \frac{r^2 n^s [\ln(2^n r_0/r)]^n}{2^n \cdot n!} dr = \sum_{n=0}^{\infty} \frac{4^n n^k}{n+3} < \infty,$$

что и требовалось доказать.

**ТЕОРЕМА 4.** *Величина  $\mathbf{D}\chi_s^{(N)}$  ограничена равномерно по  $\varepsilon$  для выпуклой трехмерной области.*

**Доказательство.** Во-первых, заметим, что из доказанной леммы здесь следует равномерная по  $\varepsilon$  и  $r_0$  ограниченность величины  $\mathbf{M}\chi_s^{(N)}$ , так как последовательность  $\{r_i\}$  для выпуклой области мажорируется такой же последовательностью, построенной в случае плоской границы по тем же реализациям изотропного направления. Далее,

$$\mathbf{M}\chi_s^2 = \mathbf{M} \sum_{n=0}^{\infty} \sum_{p=0}^{\infty} n^s p^s r_n^2 r_p^2 \leq 2\mathbf{M} \sum_{n=0}^{\infty} n^s r_n^2 \sum_{p=n}^{\infty} p^s r_n^2 \leq C' \sum_{n=0}^{\infty} n^{2s} r_n^2 < \infty.$$

Здесь было использовано условное осреднение суммы  $\sum_{p=n}^{\infty} (p - n + 1)^s r^2 s$  для фиксированного значения последовательности  $p_0, p_1, \dots, p_n$ . Теорема доказана.

ЗАМЕЧАНИЕ. Теорема 4 наверняка выполняется и в двумерном случае, так как последовательность окружностей сходится к выпуклой границе быстрее, чем последовательность сфер. Однако доказать это строго пока не удалось из-за отсутствия явного выражения плотностей  $f_n(r)$  в двумерном случае.

Конечность дисперсии рассматриваемого алгоритма в общем случае устанавливает следующая теорема.

ТЕОРЕМА 5. Дисперсия векторной оценки решения уравнения с оператором  $(\Delta - \lambda)^\gamma$  при фиксированном значении  $\varepsilon > 0$  для ограниченной области конечна.

Доказательство. Полученные оценки весов показывают, что оценка решения мажорируется конечной суммой сумм вида  $\chi_s^{(N)}$ . При  $\varepsilon > 0$  спектральный радиус субстохастического ядра, соответствующего “блужданию по сферам”, конечен и  $M\chi_s^{(N)} < \infty$  равномерно по точкам начала блужданий в ограниченной области. Далее, можно оценить величину  $M[\chi_s^{(N)}]^2$  так же, как оценивалась ранее величина  $M\chi_s^2$ . Теорема доказана.

**4. Применение алгоритма “блужданий по сферам” к решению некоторых прикладных задач.** Известно, что в теории сопротивления материалов важную роль играет матрица коэффициентов влияния  $\{\delta_{ij}\}_{i,j=1,\dots,n}$ . Здесь  $\delta_{ij}$  есть перемещение точки  $P_i$  пластины, если в точке  $P_j$  приложена единичная сила. Матрица коэффициентов влияния необходима, например, если в точках  $\{P_k\}_{k=1}^n$  расположены сосредоточенные массы и изучается колебание пластины с этими массами или если в точках  $\{P_k\}_{k=1}^n$  пластина опирается на опоры и требуется определить реакции этих опор при нагружении пластины. Рассмотрим один пример, когда перемещение упругой системы — пластины — описывается бигармоническим уравнением

$$\Delta\Delta u = \frac{1}{D}q(x, y). \quad (4.1)$$

Здесь  $u(x, y)$  — нормальный прогиб пластины;  $q(x, y)$  — интенсивность нормальной нагрузки;  $D = Eh^3/12(1 - \sigma^2)$ , где  $E$  — модуль Юнга;  $\sigma$  — постоянная Пуассона материала пластины;  $2h$  — ее толщина. Пусть край пластины свободно оперт:

$$u|_{\partial G} = 0, \quad \Delta u - \frac{1 - \sigma}{\rho} \frac{\partial u}{\partial \nu} \Big|_{\partial G} = 0. \quad (4.2)$$

Здесь  $\nu$  — внешняя нормаль к  $\partial G$ ;  $\rho$  — радиус кривизны контура  $\partial G$ . В связи с решением задачи о нахождении  $\delta_{ij}$  нас будет интересовать случай, когда  $q(P) = \delta(P - P_j)$ . Предположим еще, что  $\partial G$  — кусочно линейный контур, ограничивающий область  $G$ .

Выпишем сначала случайную оценку решения в точке  $P_0$  для случая уравнения  $\Delta\Delta u = 0$  с граничными условиями  $u|_{\partial G} = f$ ,  $\Delta u|_{\partial G} = g$ . Тогда

$$C^{(n)} = \begin{pmatrix} 1 & -\frac{1}{4} \sum_{j=1}^n r_j^2 \\ 0 & 1 \end{pmatrix},$$

поэтому случайная оценка для  $u(P_0)$  имеет вид

$$z = -\frac{q(x_N^*)}{4} \sum_{j=1}^N d^2(\xi_{j-1}) + f(x_N^*).$$

Рассмотрим теперь случай неоднородного уравнения с правой частью  $\delta(P - P_j)/D$ . Так как функция  $V(P) = r^2 \ln r / 8\pi D$ , где  $r$  — расстояние между точками  $P$  и  $P_j$ , является решением уравнения  $\Delta\Delta V(P) = \delta(P - P_j)/D$ , то решение уравнения (4.2) можно представить в виде  $u(P) = V(P) + W(P)$ , где  $\Delta\Delta W = 0$ .

Таким образом, чтобы решить приближенно задачу (4.1), (4.2) в точке  $P_i$ , можно применить следующий алгоритм. Из точки  $P_i$  моделируется процесс “блуждания по сферам”, который обрывается, скажем, в точке  $Q \in \partial G_\varepsilon$ . Тогда  $u(P_i) = W(P_i) + V(P_i)$ , где для  $W(P_i)$  статистическая оценка  $z(P_i)$  имеет вид [14]

$$z(P_i) = -\frac{|Q - P_j|^2}{8\pi D} \ln |Q - P_j| + \sum_{k=1}^{N(\varepsilon)} \frac{d^2(\xi_{k-1})}{8\pi D} [1 + \ln |Q - P_j|].$$

Отметим, что в случае распределенной нагрузки  $q(x, y)$  можно построить рандомизированный алгоритм учета правой части, основанный на представлении решения в виде свертки  $u(x_0) = u_\delta * q$ , где  $u_\delta$  — решение задачи для случая  $q = \delta(x - x_0)$ .

Следует еще отметить, что обобщение изложенного метода легко осуществляется и для случая пластин на упругом основании с известными условиями закрепления по краям, описываемыми уравнениями

$$\Delta\Delta u + \alpha\Delta u + \beta u = q(x, y).$$

Рассмотрим краевую задачу Дирихле на собственные значения для оператора Лапласа

$$\Delta u + \lambda u = 0, \quad x \in G, \quad (4.3)$$

$$u|_{\partial G} = 0, \quad x \in \partial G, \quad u \in C^2(G) \cap C(\bar{G}). \quad (4.4)$$

Как известно [15], задача (4.3), (4.4) эквивалентна задаче на собственные значения для однородного интегрального уравнения

$$u(x) = \lambda \int_G g(x, y) u(y) dy, \quad u \in C(G), \quad (4.5)$$

где  $g(x, y)$  — функции Грина задачи Дирихле для области  $G$ . Для приближенного нахождения наименьшего по модулю характеристического числа  $\lambda_1$  и соответствующей собственной функции в задаче (4.3), (4.4) можно применить метод Келлога.

Пусть функция  $\varphi^{(0)} \in L_2(G)$  не ортогональна к собственной функции, соответствующей  $\lambda_1$ . Положим

$$\varphi_p(x) = \frac{\varphi^{(p)}(x)}{\|\varphi^{(p)}\|_{L_2(G)}}, \quad \lambda_p = \frac{\|\varphi^{(p-1)}\|}{\|\varphi^{(p)}\|}, \quad p = 1, 2, \dots, \quad (4.6)$$

где  $\varphi^{(p)} = g^{(p)} \varphi^{(0)}$  — итерации функции  $\varphi^{(0)}$ . Тогда для любой функции  $\varphi^{(0)}(x) \geq 0$ ,  $\|\varphi^{(0)}\|_{L_2(G)} = 1$ , последовательность  $\{\lambda_p\}$  сходится, монотонно убывая, к  $\lambda_1$ ,  $\{\varphi_p\}$  сходится к  $\varphi$  в  $L_2(G)$  и в  $C(G)$ . Заметим, что  $\varphi^{(p)}$  удовлетворяет уравнению

$$\Delta^p \varphi^{(p-1)}(x) = (-1)^p \varphi_0(x), \quad (4.7)$$

$$\Delta^k u|_{\partial G} = 0, \quad k = 0, \dots, p-1. \quad (4.8)$$

Таким образом, алгоритм, описанный в п. 3, позволяет организовать вычисление  $\varphi_0(x_0)$  и  $\lambda_p$  в (4.6) как вычисление средних некоторых функционалов от процесса “блуждания по сферам”, а не от ветвящегося процесса [16]. Можно предложить и рандомизированный метод оценки решения уравнения (4.7) с  $p = \infty$ , что позволит вычислять сразу предельное значение  $\varphi_1$ .

### СПИСОК ЛИТЕРАТУРЫ

1. Браун Дж. Методы Монте-Карло // Современная математика для инженеров. — М.: ИЛ, 1959. — С. 275–301.
2. Елепов Г.С., Михайлов Г.А. Алгоритм “блужданий по сферам” для уравнения  $\Delta u - cu = -gg$  // Докл. АН СССР. 1973. Т. 212, № 1. С. 15–18.
3. Сипин А.С. Решение первой краевой задачи для уравнения эллиптического типа методом Монте-Карло. Методы Монте-Карло в вычислительной математике и математической физике // Труды VI Всесоюз. совещ. по методам Монте-Карло. — Новосибирск: ВЦ СО АН СССР, 1979. — Ч. 2. С. 113–119.
4. Сабельфельд К.К. Решение одной краевой задачи для метатармонического уравнения методом Монте-Карло // ЖВМиМФ. 1979. Т. 19, № 4. С. 961–969.
5. Елепов Б.С., Кронберг А.А., Михайлов Г.А., Сабельфельд К.К. Решение краевых задач методом Монте-Карло. — Новосибирск: Наука, 1979.
6. Сабельфельд К.К. О применении операторов выметания и усреднения для построения алгоритмов метода Монте-Карло // Докл. АН СССР. 1979. Т. 247, № 5. С. 1057–1060.
7. Хасьминский Р.З. Вероятностное представление решений некоторых дифференциальных уравнений // Труды VI Всесоюз. совещ. по теории вероятностей и математической статистике. — Вильнюс: ГИПНА, 1960. — С. 177–183.
8. Гельфанд И.М., Фролов А.С., Ченцов Н.Н. Вычисление континуальных интегралов методом Монте-Карло // Изв. вузов. Сер. Математика. 1958. № 5 (6). С. 32–45.
9. Елепов Б.С., Михайлов Г.А. О решении задачи Дирихле для уравнения  $\Delta u - cu = -g$  моделированием “блужданий по сферам” // ЖВМиМФ. 1969. Т. 9, № 3. С. 647–654.
10. Кронберг А.А. К решению двух краевых задач методом Монте-Карло // ЖВМиМФ. 1976. Т. 16, № 1. С. 162–169.
11. Михайлов Г.А. Дисперсия векторных алгоритмов метода Монте-Карло // Докл. АН СССР. 1980. Т. 253, № 5. С. 1047–1050.
12. Сабельфельд К.К., Елепов Б.С. Решение одной краевой задачи для полигармонического уравнения методом Монте-Карло // Численные методы механики сплошной среды. 1976. Т. 7, № 7. С. 69–83.
13. Гантмахер Ф.Р. Теория матриц. — М.: Наука, 1967.
14. Сабельфельд К.К., Ультан А.Е. Расчет матрицы коэффициентов влияния для пластины методом Монте-Карло. Методы Монте-Карло в вычислительной математике и математической физике // Труды VI Всесоюз. совещ. 1979. Ч. 2. С. 120–124.
15. Владимиров В.С. Уравнения математической физики. — М.: Наука, 1973.
16. Кронберг А.А. Об одном сочетании метода Келлога с методом Монте-Карло. Методы Монте-Карло в вычислительной математике и математической физике / Под ред. Г. И. Марчука. — Новосибирск: ВЦ СО АН СССР, 1974. — С. 136–142.

## ПРИБЛИЖЕННОЕ ВЫЧИСЛЕНИЕ КОЭФФИЦИЕНТОВ УГЛОВОГО РАСПРЕДЕЛЕНИЯ ДЛЯ НЕСТАЦИОНАРНОГО ТЕПЛОВОГО ИЗЛУЧЕНИЯ С ПОВЕРХНОСТИ ВРАЩЕНИЯ<sup>\*)</sup>

Интегральное уравнение Ю.Н. Бабаева [1]

$$u(P, t) = \int_{S(P)} u\left(Q, t - \frac{\rho(P, Q)}{c}\right) K(P, Q) dS_Q - q(P, t), \quad (1)$$

$$K(P, Q) = \frac{1}{\pi} \frac{\cos(\vec{n}_P, \vec{PQ}) \cos(\vec{n}_Q, \vec{QP})}{\rho^2(P, Q)}, \quad (2)$$

описывает нестационарный перенос энергии излучением через область с пренебрежимо малым взаимодействием излучения и вещества. Здесь  $u$  — плотность потока энергии, излучаемой поверхностью  $S$  внутрь полости;  $q$  — заданная функция, плотность потока в окружающее пространство. Ядро  $K(P, Q)$  соответствует излучению с поверхности по закону Ламберта,  $\vec{n}$  — внутренняя по отношению к полости нормаль к поверхности  $S$ . Учитывается запаздывание на время прохождения расстояния  $\rho(P, Q)$ ,  $c$  — скорость света. Энергия в точку  $P, Q \in S$  собирается с части  $S(P)$  поверхности  $S$ , не затененной для точки  $P$  самой поверхностью  $S$ : отрезок  $PQ$  не пересекает поверхность  $S$ ; в этом случае о точках  $P, Q$  мы будем говорить, что они *видят друг друга*, а область интегрирования  $S(P)$  будем называть *областью видимости* для точки  $P$ .

Наряду с уравнением лучистой теплопроводности уравнение (1) является важным приближением уравнения переноса излучения [2], позволяющим оценить кинетические эффекты и рассчитать энергоперенос излучением через область малой оптической толщины.

Рассмотрим двумерную задачу, решение которой обладает симметрией вращения относительно некоторой оси. Пусть поверхность  $S$  состоит из нескольких односвязных аналитических поверхностей вращения  $S_\alpha$ ,  $S = \bigcup_\alpha S_\alpha$ , граничащих по окружностям  $K_{\alpha\alpha'}$ .

Для численного решения задачи введем на поверхности  $S$  счетную сетку так, как это делается в стационарных расчетах зональным методом [3], включая в число сеточных линий окружности  $K_{\alpha\alpha'}$ , чтобы исключить разрывы решения внутри счетных ячеек [1].

Счетными областями  $S_{ik}$  нам будут служить части  $S_i$  заданных аналитических поверхностей, сами являющиеся поверхностями вращения, дополнительно

<sup>\*)</sup> ВАНТ. Сер. Методики и программы численного решения задач математической физики. 1983. Вып. 1 (12). С. 32–34 (совм. с Ю.А. Деметьевым, Р.Ф. Машининым, В.Ф. Мироновой).

рассеченные полуплоскостями  $\varphi = \text{const}$  ( $\varphi$  — угол поворота вокруг оси):

$$S = \bigcup_i S_i, \quad S_i = \bigcup_k S_{ik}, \quad S = \bigcup_{i,k} S_{ik}.$$

Уравнение (1) с учетом заданной симметрии вращения аппроксимируется системой алгебраических уравнений

$$S_i u_i(t) = \sum_{j,k} \alpha_{ijk} u(t - l_{ijk}) - S_i q_i(t), \quad (3)$$

$$\alpha_{ijk} = \int_{S_i} \int_{S_{jk}(P)} K(P, Q) dS_Q dS_P, \quad P \in S_i, \quad Q \in S_{jk}, \quad (4)$$

$$l_{ijk} = \frac{1}{\alpha_{ijk}} \int_{S_i} \int_{S_{jk}(P)} \rho(P, Q) K(P, Q) dS_Q dS_P. \quad (5)$$

Коэффициенты  $\alpha_{ijk}, l_{ijk}$  удовлетворяют соотношениям

$$\sum_{jk} \alpha_{ijk} = S_i, \quad \sum_{ijk} \alpha_{ijk} l_{ijk} = 4V, \quad (6)$$

$V$  — объем полости. Отметим также симметрию матрицы коэффициентов  $\alpha_{ij}$  взаимодействия любых двух зон  $S_i, S_j$  на поверхности  $S$ :

$$\alpha_{ij} = \sum_k \alpha_{ijk}, \quad \alpha_{ij} = \alpha_{ji}. \quad (7)$$

Коэффициенты  $\alpha_{ijk}, l_{ijk}$  концентрируют всю геометрическую информацию о расчете:

$\alpha_{ijk}/S_i$  — интегральные угловые коэффициенты излучения;

$l_{ijk}$  — средние расстояния между зонами на поверхности  $S$  задания граничных условий  $q(t)$ .

Вычисление коэффициентов взаимодействия сеточных элементов поверхности отделяется в самостоятельную задачу от собственно расчета переноса излучения. В этой статье описывается методика приближенного вычисления с использованием ЭВМ коэффициентов излучения для поверхностей вращения, опирающаяся, по возможности, на аналитические расчеты. Предлагаемый алгоритм уточняет и обобщает алгоритм [4].

В качестве переменных интегрирования в уравнениях (4), (5) выберем длины образующих  $r_i, r_j$  сеточных областей  $S_i, S_{jk}$  и углы поворота  $\varphi_i, \varphi_j$  соответственно. В силу симметрии вращения интегрирование по одному из углов  $\varphi_i$  сводится к умножению на  $2\pi$  и предполагается, что точка  $P$  принадлежит полуплоскости  $\varphi = 0$ . В практически важном случае, когда в качестве  $S_\alpha$  используются поверхности вращения отрезков прямых, удастся провести аналитическое интегрирование еще по одному из оставшихся параметров. По остальным параметрам, за исключением отдельных случаев [5], придется интегрировать каким-либо приближенным способом.

Необходимо еще определить область видимости  $S_{jk}(P)$  для точек  $P, P \in S_i$ . Это делается с помощью следующих двух алгоритмов (п. 1, 2).

1. Для точки  $P \in S_i$  определим аналитически границы  $\varphi_{j,m}\varphi_{j,m+1}$  дуг видимости  $\varphi_{j,m} \leq \varphi_j \leq \varphi_{j,m+1}$  на окружности  $Q(\varphi)$  вращения точки  $Q$ ,  $Q \in S_j$  вокруг оси симметрии.

1.1. Сначала определим границы дуг видимости с учетом возможного затенения от единственной поверхности вращения  $S_\alpha$ , ограниченной окружностями  $K_{1\alpha}, K_{2\alpha}$ . Решая задачу о пересечении отрезка  $PQ$  и поверхности  $S_\alpha$ , определим, видит ли точка  $P$  точки  $Q(0), Q(\pi)$ .

1.1.1. Если точка  $P$  видит обе точки  $Q(0), Q(\pi)$ , то возможны два случая.

Либо точка  $P$  видит всю окружность  $Q(\varphi)$ , если отрезок  $PQ(\pi)$  не пересекает круги  $K_{1\alpha}, K_{2\alpha}$ . В этом случае, учитывая симметрию относительно плоскости  $\varphi = 0$ , необходимо интегрировать по  $\varphi_j$  в пределах  $0 \leq \varphi_j \leq \pi$ .

Либо имеются две дуги видимости  $0 \leq \varphi_j \leq \varphi_{1\alpha}, \varphi_{2\alpha} \leq \varphi_j \leq \pi$ , если отрезок  $PQ(\pi)$  пересекает круги  $K_{1\alpha}, K_{2\alpha}$ . Величины  $\varphi_{1\alpha}, \varphi_{2\alpha}$  определяются, в свою очередь, из следующих трех задач: о нахождении точки  $M_0$  касания отрезка  $PQ(\varphi)$  и поверхности  $S_\alpha$ , о нахождении общих точек  $M_1, M_2$  отрезка  $PQ(\varphi)$  и каждой из окружностей  $K_{1\alpha}, K_{2\alpha}$ . Значение  $\varphi_{1\alpha}$  является минимальным, а значение  $\varphi_{2\alpha}$  — максимальным значением угла поворота, соответствующего точкам  $M_0, M_1, M_2$ ,

$$\varphi_{1\alpha} = \min_n \varphi(M_n), \quad \varphi_{2\alpha} = \max \varphi(M_n),$$

$$n = 0, 1, 2, \quad M_0 \in S_\alpha, \quad M_1 \in K_{1\alpha}, \quad M_2 \in K_{2\alpha}.$$

1.1.2. Если точка  $P$  видит точку  $Q(0)$  и не видит точку  $Q(\pi)$ , то имеется лишь одна дуга видимости  $0 \leq \varphi_j \leq \varphi_{2\alpha}$ .

1.1.3. Если точка  $P$  не видит точку  $Q(0)$  и видит точку  $Q(\pi)$ , то имеется лишь одна дуга видимости  $\varphi_{2\alpha} \leq \varphi_j \leq \pi$ .

1.1.4. Наконец, если  $P$  не видит обе точки  $Q(0), Q(\pi)$ , то точка  $P$  не видит все точки окружности  $Q(\varphi)$ .

1.2. Обобщая случаи 1.1.1–1.1.4, можно сказать, что всегда имеются две дуги видимости  $0 \leq \varphi_j \leq \varphi_{1\alpha}, \varphi_{2\alpha} \leq \varphi_j \leq \pi$ , но в частных случаях допускаются предельные значения  $\varphi_{1\alpha} = 0, \varphi_{2\alpha} = \pi$ , при которых дуги могут иметь нулевую длину. Если учитывать возможное затенение от всех участков аналитичности  $S_\alpha$ , составляющих поверхность  $S$ , то область видимости  $Q_P$  на окружности  $Q(\varphi)$  для точки  $P$  может состоять из нескольких дуг видимости, определяемых как пересечение дуг видимости от каждой из поверхностей  $S_\alpha$ ,

$$Q_P = \bigcap_{\alpha} ((0, \varphi_{1\alpha}) \cup (\varphi_{2\alpha}, \pi)). \quad (8)$$

1.3. Конкретное применение алгоритма 1.1–1.2 для определения границ области интегрирования  $S_{jk}(P)$ , входящей в формулы (4), (5), зависит от метода численного интегрирования. Так, при интегрировании по  $r_i, r_j$  с помощью формулы центральных прямоугольников точку  $P$  следует выбирать в середине образующей для каждого участка разбиения, а дуги видимости следует искать на средней линии  $Q(\varphi)$  соответствующего участка разбиения  $\Delta r_j$ .

2. Время вычисления коэффициентов  $\alpha_{ijk}, l_{ijk}$  с помощью ЭВМ пропорционально числу поверхностей  $S_\alpha$ , если для определения области интегриро-

вания  $S_{jk}(P)$  пользоваться алгоритмом 1.1–1.2. Поэтому целесообразно составить список тех поверхностей  $S_{m\alpha}$ , которые мешают видимости каких-либо точек  $P, Q$ ,  $P \in S_\alpha$ ,  $Q \in S_\alpha$ , пересекая отрезок  $PQ$ .

Разработан алгоритм автоматического определения мешающих видимости поверхностей вращения  $S_{m\alpha}$ , составляющих часть от поверхностей  $S_\alpha$  или совпадающих с ними,  $S_{m\alpha} \subset S_\alpha$ . Алгоритм опирается на упорядоченное описание информации о  $S_\alpha$ , фиксирующее направление обхода контура образующей поверхности  $S$  в полуплоскости  $\varphi = 0$ . Вместо заданных аналитически образующих поверхностей  $S_\alpha$  рассматривается ломаная линия  $R_n$ , аппроксимирующая образующую поверхности  $S$  в полуплоскости  $\varphi = 0$  с шагом по длине образующей  $r$ , совпадающим с шагом приближенного интегрирования по  $r_i, r_j$  при вычислении интегралов (4), (5).

Границы мешающих видимости участков поверхности  $S$  некорректно зависят от погрешности в определении узлов ломаной  $R_n$ . Малейшая погрешность в определении координат узлов ломаной, лежащих на одной прямой, может привести к образованию сколь угодно больших участков, мешающих видимости. Указанная некорректность ликвидируется с помощью анализа величины второй разности в предположении ограниченности радиуса кривизны для непрямых участков образующей.

**3.** Предлагаемый алгоритм реализован в программе для вычисления коэффициентов  $\alpha_{ijk}, l_{ijk}$  в случае, когда в качестве  $S_\alpha$  используются поверхности второго порядка, получаемые вращением отрезков прямых и дуг окружности с центром на оси вращения. Расчеты показали хорошую работоспособность алгоритма. В частности, была рассчитана задача, приведенная в работе [4]. В окрестности точек резкого излома поверхности  $S$  могут складываться условия, когда подынтегральная функция  $K(P, Q)$  имеет интегрируемую особенность. Без привлечения специальных алгоритмов интегрирования (например, приближенного интегрирования непосредственно по телесному углу) коэффициенты взаимодействия счетных ячеек, прилегающих к вершинам углов на образующей поверхности  $S$ , могут досчитываться из контрольных соотношений (6).

### СПИСОК ЛИТЕРАТУРЫ

1. Бабаев Ю.Н., Баженов С.В., Дементьев Ю.А. Об особенностях решения одного интегрального уравнения переноса излучения // Вопросы атомной науки и техники. 1978. Вып. 1 (1). С. 7–9.
2. Зельдович Я.Б., Райзер Ю.П. Физика ударных волн и высокотемпературных гидродинамических явлений. — М.: Наука, 1966.
3. Кутателадзе С.С. Основы теории теплообмена. — М: Атомиздат, 1979.
4. Ченцов Н.Н., Думкина Л.В., Шойдина Н.С. Вычисление угловых коэффициентов излучения в осесимметрической геометрии // Инженерно-физический журнал. 1978. Т. 3, № 2. С. 306–312.
5. Зигель Р., Хауэлл Дж. Теплообмен излучением. — М: Мир, 1975.

## ГЕОМЕТРИЧЕСКАЯ ХАРАКТЕРИЗАЦИЯ СФЕРИЧЕСКИХ МАТОВЫХ ОТРАЖАТЕЛЕЙ<sup>\*)</sup>

Работа посвящена геометрической оптике матовых зеркал, рассеивающих излучение по закону Ламберта. В принятой здесь схематизации единственной характеристикой излучения является его интенсивность, а характеристикой отражающей поверхности — ее альbedo. Взаимодействием луча и окружающей среды между актами отражения мы пренебрегаем. В такой геометрической теории рассеяния хорошо известно следующее свойство сферической поверхности.

Пусть  $S$  — абсолютно черная сферическая поверхность. Если некоторый ее участок  $\Delta S_1$  излучает внутрь по закону Ламберта, т.е. с индикатрисой

$$dE = \frac{E}{\pi} \cos \theta \sin \theta d\theta d\varphi$$

(где  $\theta$  — полярный угол,  $\varphi$  — азимутальный угол), то освещенность остальной поверхности во всех точках одинакова.

Это свойство делает матовый серый сферический экран, рассеивающий по закону Ламберта, идеальным гомогенизатором излучения: интенсивность отразившегося от экрана “света” постоянна внутри сферы и изотропна по всем направлениям. Мы доказываем, что это свойство сферической поверхности является характеристическим.

**ТЕОРЕМА.** Пусть  $S$  — абсолютно черная кусочно гладкая замкнутая ориентированная поверхность без края, обладающая следующим свойством: любой ее участок  $\Delta S_1$ , излучающий по закону Ламберта, равномерно (т.е. с одинаковой интенсивностью, отличной от нуля) освещает остальную поверхность (точнее, ее сторону, отвечающую положительному направлению нормали). Тогда поверхность  $S$  — сферическая полость.

**Доказательство.** Поверхностный источник  $dS_1$  в точке  $\vec{r}_1$  с нормалью  $\vec{n}_1$ , и с ламбертовой индикатрисой излучения приносит на участок  $dS_2$  в точке  $\vec{r}_2$  с нормалью  $\vec{n}_2$  “световой” поток, пропорциональный коэффициенту влияния

$$dE = \frac{\cos \psi_2 \cos \psi_1}{|\vec{r}_{12}|^2} dS_1 dS_2, \quad (1)$$

где вектор  $\vec{r}_{ij} = \vec{r}_i - \vec{r}_j$ ,  $|\vec{r}_{ij}| \cos \psi_i = (\vec{n}_i, \vec{r}_{ij})$  [1]. Таким образом, поверхность  $S$  по условию теоремы обладает свойством: для любых двух точек  $\vec{r}_1, \vec{r}_2 \in S$  коэффициент влияния

$$\frac{\cos \psi_2 \cos \psi_1}{|\vec{r}_{12}|^2} = \frac{(\vec{n}_1, \vec{r}_{12})(\vec{n}_2, \vec{r}_{21})}{|\vec{r}_{12}|^4} = \text{const} > 0. \quad (2)$$

<sup>\*)</sup> ВАНТ. Сер. Методики и программы численного решения задач математической физики. 1983. В. 1 (12). С. 37–40 (совм. с О.Б. Феодоритовой).

Проведем сперва краткое доказательство в предположении, что  $S$  дважды непрерывно дифференцируема.

Рассмотрим две близкие точки  $\vec{r}_1$  и  $\vec{r}_2 = \vec{r}_1 + \Delta\vec{r}$ . Для смещений по главному направлению  $d\vec{r} = d\tau \cdot \vec{t}_k$  имеем формулу Родрига

$$d\vec{n} = -\varkappa_k d\vec{r} = -\varkappa_k d\tau \cdot \vec{t}_k, \quad (3)$$

т.е. обе нормали  $\vec{n}_1$  и  $\vec{n}_2$  лежат в одной плоскости с точностью до малых высшего порядка. С другой стороны, для плоской кривой — сечения  $S$  плоскостью, проходящей через  $\vec{n}_1$  и  $\vec{t}_k$  имеем с точностью до малых следующего порядка

$$\Delta\vec{r} = \Delta\tau \cdot \vec{t}_k + \frac{1}{2}\varkappa_k(\Delta\tau)^2\vec{n}_1, \quad \vec{n}_2 = \vec{n}_1 - \varkappa_k\Delta\tau \cdot \vec{t}_k.$$

Отсюда

$$(\vec{n}_1, \vec{r}_{12}) \cdot (\vec{n}_2, \vec{r}_{12}) \approx \varkappa_k^2(\Delta\tau)^4/4$$

с точностью до малых пятого порядка. Следовательно, в точке  $\vec{r}_1$  главные кривизны (отвечающие максимуму и минимуму кривизны сечений) должны быть одинаковы. Но тогда это точка округления: все направления — главные, все кривизны нормальных сечений — одинаковые. А как известно [2], поверхность без края, сплошь состоящая из омбилических точек, — сфера.

Идею этого доказательства удастся перенести на общий случай. Сначала докажем вспомогательное предположение, имеющее самостоятельный интерес.

**ЛЕММА.** Пусть в вершинах  $A$  и  $C$  прямоугольного треугольника  $AOC$  с прямым углом  $O$  расположены перпендикулярно к соответствующим катетам площадки  $dS_A$  и  $dS_C$ . Тогда при фиксированной длине катета  $|AO| = R$  выражение (3) при приближении точки  $C$  из бесконечности к  $O$  сначала монотонно возрастает, достигает максимума для  $\psi_A = \pi/6$ ,  $\psi_C = \pi/3$ ,  $|OC| = R/\sqrt{3}$ , а затем монотонно убывает до нуля при  $C = O$ .

**СЛЕДСТВИЕ.** Освещенность площадки  $dS_C$  от источника  $dS_A$ , возникающая на расстоянии  $|OC| = R$ , совпадает с аналогичной освещенностью на расстоянии  $|OC| = \rho$ ,  $\rho < R/\sqrt{3}$ , и только с ней.

Доказательство леммы. В этом случае  $|r_{12}| = R/\cos\psi$ ,  $\psi_A = \psi$ ,  $\psi_C = -\psi = \pi/2$ , так что по выражению (2) коэффициент влияния

$$K(\psi) = R^{-2} \sin\psi \cos^3\psi.$$

Он достигает максимума при  $K'(\psi) = 0$ , т.е. при  $3\tg^2\psi_0 = 1$ ,  $\psi_0 = \pi/6$ .

Приступим к доказательству самой теоремы. Из выражения (2) немедленно вытекает: 1) расстояния между точками  $S$  ограничены, т.е. вся поверхность ограничена; 2) поверхность  $S$  лежит по одну сторону от любой своей касательной плоскости, т.е. поверхность выпуклая; 3) направления нормалей в близких точках близки, т.е. поверхность непрерывно дифференцируема, и у нее нет ни ребер, ни точек заострения.

Так как поверхность  $S$  замкнута и, как доказано, ограничена, то на ней най-

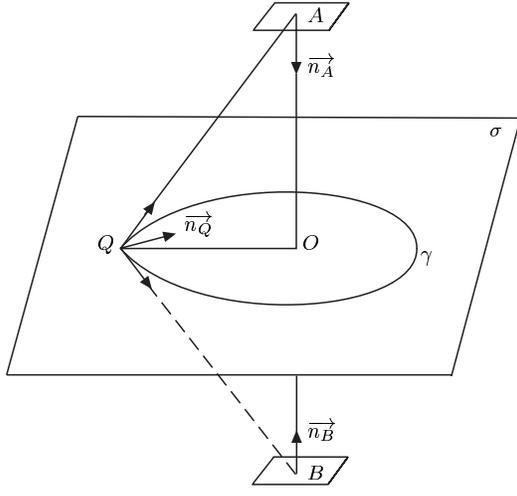


Рис. 1

должны давать одинаковую освещенность. Так как  $Q$  равноудалена от  $A$  и  $B$ , то

$$r = |r_{AQ}| = |r_{BQ}|,$$

$$\cos \psi_{AQ} = \cos \psi_{BQ} = \frac{R}{r}.$$

Отсюда

$$\cos \psi_{QA} = \cos \psi_{QB},$$

т.е. нормаль к  $dS_Q$  должна образовывать равные углы с  $r_{QA}$  и с  $r_{QB}$  (см. рис. 1). Отсюда она должна лежать в плоскости  $\sigma$ .

Рассмотрим теперь на  $\gamma$  точки  $C$  и  $D$ , находящиеся от  $O$  на максимальном и минимальном удалении. Так как  $dS_A$  должно создавать одинаковую освещенность и на  $dS_B$ , и на  $dS_C$ , и на  $dS_D$ , то по лемме, либо  $|OC| = |OD| = R$ , либо  $|OC| = R$ ,  $|OD| = \rho < R\sqrt{3}$ , либо  $|OC| = |OD| = \rho$ . Последний случай невозможен, так как тогда вся кривая  $\gamma$  — окружность радиуса  $\rho$ . Но тогда для любых

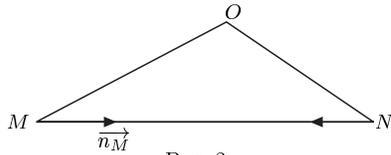


Рис. 2

двух ее противоположных точек  $M$  и  $N$  (рис. 2) нормали лежат на прямой  $MN$ , т.е. участки  $dS_M$  и  $dS_N$  освещают друг друга с интенсивностью  $(2\rho)^{-2} > (2R)^{-2}$ , отличной от таковой для  $dS_A$  и  $dS_B$ . Обратимся к исследованию второго случая. Пусть точки  $M, N \in \gamma$  максимально удалены друг от друга. Тогда плоские нормали к  $\gamma$  в этих точках лежат на  $MN$ . Поскольку на  $\sigma$  они, по доказанному, совпадают с  $n_M$  и  $n_N$ , то площадки  $dS_M$  и  $dS_N$  освещают друг друга с интенсивностью  $|MN|^{-2}$ . Поэтому  $|MN| = 2R$ . Так как  $|OM|, |ON| \leq |OC| = R$ , то по неравенству треугольника  $|OM| = |ON| = R$  и точка  $O$  — середина отрезка  $MN$  (см. рис. 2). Раз точки  $M$  и  $N$  находятся на максимально возможном для поверхности  $S$  удалении, то к ним применимы выводы, полученные для  $A$  и  $B$ . В частности, для точек  $K$

дуются две точки  $A$  и  $B$  (рис. 1), лежащие на максимальном расстоянии:  $|AB| = 2R$ . Как известно, для этих точек касательные плоскости ортогональны прямой  $AB$ , откуда

$$\frac{\cos \psi_A \cos \psi_B}{|AB|^2} = \frac{1}{4R^2}.$$

Пусть  $O$  — середина отрезка  $AB$ . Проведем через нее плоскость  $\sigma$ , перпендикулярную  $AB$ . Она пересечет выпуклую поверхность  $S$  по выпуклой кривой  $\gamma$ , окружающей  $O$ . Для любой точки  $Q \in \gamma$  на касательной площадке  $dS_Q$  площадки  $dS_A$  и  $dS_B$

сечения  $S$  плоскостью  $\tau$ , проходящей через  $O$  перпендикулярно  $MN$ , все нормали  $n_k$  лежат в этой плоскости (рис. 3). Пусть  $K$  и  $L$  — две точки пересечения  $\gamma \cap \tau = S \cap \sigma \cap \tau$ . Нормали  $n_K$  и  $n_L$  лежат как в  $\sigma$ , так и в  $\tau$ , т.е. лежат на прямой  $KL$ . Отсюда, по постоянству коэффициента влияния (2),  $|KL| = 2R$ ,  $|OK| = |OL| = R$ . Мы показали, что в выпуклую кривую  $\gamma$  можно вписать квадрат  $KMLN$  с центром  $O$  и полудиagonalю  $R$ . Все точки периметра квадрата удалены от  $O$  на расстояние, не меньшее  $R/\sqrt{2}$ . Точки кривой  $\gamma$  не лежат внутри квадрата, поэтому на  $\gamma$  ни одной точки  $D$  с расстоянием  $|OD| = \rho < R\sqrt{3}$  нет. Значит, второй случай не реализуется, т.е. минимальное удаление  $|OD| = |OC| = R$ , и все точки  $\gamma$  лежат на окружности радиуса  $R$ . Итак, “экватор” нашей поверхности — окружность радиуса  $R$ . Применяя этот вывод вместо полюсов  $A$  и  $B$  к любой паре диаметрально противоположных точек экватора, убеждаемся, что и все “меридианы” являются окружностями радиуса  $R$  с центром в  $O$ . Следовательно,  $S$  — сфера.

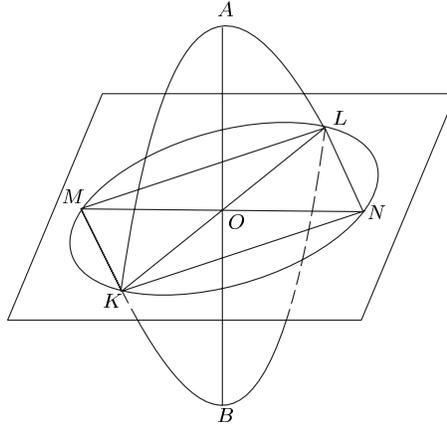


Рис. 3

Доказанная теорема о выравнивании освещенности на втором рассеянии справедлива, если в сферической полости нет мешающих тел. Однако если мешающие образования малы, освещенность оказывается довольно близка к равномерной (поскольку мы рассматриваем процесс по числу актов рассеяния, а не в реальном времени, речь всюду идет о суммарных по времени эффектах).

Рассмотрим простейшую модель, когда внутрь рассеивающего с альбедо  $\alpha$  сферического экрана помещено ядро — концентрическое сферическое абсолютно черное тело. Такая простейшая модель лучистого переноса может быть рассчитана аналитически. В свое время она была использована нами для тестирования погрешностей при расчете матриц влияния в методике расчета переноса в работе [3].

Предположим для определенности, что первоначально излучение было сконцентрировано на экране. Интегрируя излучение из точки на поверхности экрана по конусу, опоясывающему ядро, получаем  $e_{\text{погл}}^{(n+1)} = \gamma^2 E_{\text{расc}}^{(n)}$ , где  $n$  — номер рассеяния,  $\gamma = rR^{-1}$  — отношение радиуса сферы к радиусу экрана. Суммируя прогрессию по числу рассеяний, получаем, что суммарная энергия, поглощенная черной сферой, равна

$$e_{\text{погл}} = \frac{\gamma^2}{1 - \alpha(1 - \gamma^2)} E_{\text{нач}}. \tag{4}$$

Коэффициент  $\rho(\alpha, \gamma) = \gamma^2 [1 - \alpha(1 - \gamma^2)]^{-1}$  при  $E_{\text{нач}}$  строго меньше единицы (и равен ей лишь при  $\gamma = 1$  или  $\alpha = 1$ ). Если считать, что первоначальное излучение было индуцировано на ядре, то

$$e_{\text{погл}} = \alpha \rho(\alpha, \gamma) E_{\text{нач}}. \tag{5}$$

Зависимость этого коэффициента  $\alpha \rho(\alpha, \gamma)$  от  $\alpha$  и  $\gamma$  приведена ниже:

$\alpha$	$\gamma$								
	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
0,5	0,01	0,038	0,083	0,138	0,2	0,265	0,329	0,39	0,448
0,6	0,015	0,057	0,119	0,194	0,272	0,351	0,424	0,490	0,549
0,7	0,023	0,085	0,174	0,272	0,368	0,457	0,534	0,599	0,654
0,8	0,038	0,138	0,265	0,390	0,5	0,590	0,662	0,719	0,764
0,9	0,082	0,265	0,448	0,590	0,692	0,764	0,815	0,852	0,879
0,99	0,497	0,798	0,899	0,941	0,961	0,977	0,980	0,984	0,988
0,95	0,160	0,432	0,631	0,752	0,826	0,872	0,903	0,924	0,939
1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0

Как видим, при малых  $r$  коэффициент использования зависит от  $r$  квадратически. В общем случае, когда внутренняя сфера серая, с альбедо  $\beta$ , приходим вместо (5) к более сложному выражению

$$e_{\text{погл}}(\alpha, \beta, \gamma) = (1 - \beta)\alpha\rho(\alpha, \gamma)[1 - \alpha\beta\rho(\alpha, \gamma)]^{-1}E_{\text{нач}}. \quad (6)$$

Аналогична формула и для случая, когда первоначальное излучение индуцируется на экране:

$$e_{\text{погл}}(\alpha, \beta, \gamma) = (1 - \beta)\rho(\alpha, \gamma)[1 - \alpha\beta\rho(\alpha, \gamma)]^{-1}E_{\text{нач}}. \quad (7)$$

Средняя длина пробега между рассеяниями на поверхности экрана также вычисляется квадратурой

$$l = \frac{4}{3}R(1 - \gamma^3), \quad (8)$$

что дает средний путь до поглощения на черном ядре при условии, что поглощение произойдет именно на ядре,

$$L = \frac{4}{3}R\gamma^3 + \frac{4}{3}R(1 - \gamma^3)\alpha(1 - \gamma^2)[1 - \alpha(1 - \gamma^2)]^{-1}. \quad (9)$$

Оценка распределения принятой черной сферой рассеянной энергии требует знания угловых коэффициентов влияния [3]. Нами был предпринят расчет модельного примера с  $\gamma = 0,179$ , в котором излучение индуцировалось равномерно (по площади) в приэкваториальной "половине" черного ядра, где полярный угол  $\pi/3 < \theta < 2\pi/3$ . Для альбедо  $\alpha = 0,5$  максимальный возврат был на экваторе — плотность возвратившейся энергии  $\Delta E/\Delta S = 1,5 \cdot 10^{-2}$ , где плотность переизлученной энергии взята за единицу, а минимальный возврат  $1,18 \cdot 10^{-2}$  был на полюсах. Это дает неравномерность вложения порядка 23 %, порождаемую в основном неравномерностью однократно рассеянного излучения.

Поскольку для выравнивания надо греть полюса за счет экватора, то мы учли результат леммы о максимуме функции влияния и сплющили экран у полюсов. Была рассчитана модельная конфигурация с экраном в форме сплющенного эллипсоида вращения того же объема, что и у исходного экрана,  $b = a\sqrt{3}$ ,  $ab^2 = R^3$  и тем же сферическим ядром. Расчет привел к цифрам  $1,34 \cdot 10^{-2}$  на экваторе и  $1,2 \cdot 10^{-2}$  на полюсах черного ядра. Таким образом, неравномерность оказалась лишь 11 %, что вдвое меньше, чем в сферическом случае, при незначительном уменьшении интеграла возвратившейся энергии 0,0286 вместо 0,0309.

Для проведения расчетов пришлось модернизировать программу, описанную в работе [3], введя в нее в качестве элементарных и поверхности эллипсоидов вращения. Площади эллипсоидальных поясов вычислялись по квадратурам Гюйгенса, а формулы для размера тени от конуса из работы [3] были заменены не более общую формулу для размера тени от поверхности вращения второго порядка. По модернизированной программе были проведены также расчеты конфигураций с экраном в форме вытянутого эллипсоида вращения, давшие в описанной выше модельной задаче большую асимметрию, чем при сферическом экране.

Круг задач физики, описываемых принятой в работе геометро-оптической схемой, довольно обширен. Однако во многих вопросах, например, в физике высокотемпературных явлений, где ее выдвинул Я.Б. Зельдович, она является лишь первым грубым приближением. Так, в частности, будет, когда собственное излучение звезды заперто ее непрозрачным верхним слоем, [4]. В более сложной модели рассматривается дополнительно непрозрачное ядро звезды.

#### СПИСОК ЛИТЕРАТУРЫ

1. Исаченко В.П., Осипова В.А., Сухомел А.С. Теплопередача. — М.-Л.: Энергия, 1969.
2. Фавар Ж. Курс локальной дифференциальной геометрии. — М.: ИЛ, 1960.
3. Ченцов Н.Н., Думкина Г.В., Шойдина Н.С. Вычисление угловых коэффициентов в осесимметричной геометрии // Инженерно-физический журнал. 1978. Т. 34, № 2.
4. Зельдович Я.Б., Блииников С.И., Шакура Н.И. Физические основы строения и эволюции звезд. — М.: МГУ. 1981.

## ПРИБЛИЖЕННЫЕ РЕШЕНИЯ В БАЛЛИСТИЧЕСКОЙ ЗАДАЧЕ ЛАГРАНЖА И ЕЕ ОБОБЩЕНИЯХ \*)

### Введение

Одномерные газодинамические течения, в которых состояние среды и конфигурация системы зависят только от одной пространственной декартовой координаты  $x$  и еще от времени  $t$ , удобно описывать в терминах движения объемов невязкого газа в длинных цилиндрических трубах, например мысленно вырезанных из системы [1]. Итак, пусть дана бесконечная пустая цилиндрическая труба сечения  $\mathfrak{S}$ , в которой могут двигаться без трения несжимаемые поршни, разделенные объемами газа или пустотой. Предположим вслед за Лагранжем [2], что один из поршней (“казенник”) закреплен неподвижно, а “камера” между ним и покоящимся вторым поршнем массы  $M$  (“ядром”, по терминологии Лагранжа и

\*) Ин-т прикл. матем. им. М.В. Келдыша АН СССР, М., 1988 /Препринт № 58. (совм. с А.В. Забродиным).

Пуассона) заполнена в начальный момент  $t_0$  неподвижным однородным сильно сжатым газом массы  $m$  и удельной энергии  $E_0$  с показателем адиабаты  $\gamma$ . Требуется определить, до какой финальной скорости разгонит “ядро” расширяющийся газ. Если перед поршнем — пустота и  $m \ll M$ , то по закону сохранения энергии  $U_\infty = \sqrt{2E/M}$ , где  $E = mE_0$  — полная энергия газа в начальный момент. Однако в общем случае, когда  $m$  сравнимо с  $M$ , вычисление кинетической энергии газа и величины  $U_\infty$  становится достаточно сложной задачей [3]; см. также [4].

Заменяв неподвижный “казенник” на покоящийся в начальный момент  $t_0$  невесомый поршень, мы переходим от весьма упрощенной постановки задачи внутренней баллистики к столь же упрощенной постановке задачи о метании пластины [5, 6]. Сходные конфигурации, но уже с тремя и более поршнями возникают при моделировании слоистых систем из [7] с кумуляцией энергии. Далее, в задаче Лагранжа интересен вопрос: какую долю от энергии  $E = mE_0$  отберет поршень к моменту  $t$ , или, скажем, пройдя путь  $l$ ? Последняя задача допускает сферическое (а также цилиндрическое) обобщение. Заменяв плоскость “казенника” на недеформируемую поверхность сферической полости, а плоский поршень — сферической оболочкой из несжимаемой жидкости, мы переходим в рамки задачи о сжатии оболочки [8].

В принятой Лагранжем постановке баллистической задачи ударных волн не возникает, и решение во все времена остается непрерывным и, следовательно, изэнтропическим, [3]. Однако аналитическое решение этой задачи оказывается сложным. В частности, оно зависит от того, сколько отражений испытает фронт волны разрежения, возникающей, когда поршень трогается с места. Все эти замечания сохраняют силу и для указанных выше обобщений задачи Лагранжа. Поэтому представляют определенный интерес компактные приближенные формулы, позволяющие свободно ориентироваться в решении. Такая потребность возникает, например, при проектировании течений с нужными свойствами [9]. При этом в плоских задачах близость движения поршней к финальному режиму можно, следуя [3], оценить на примерах автомобильных движений газа с поршнями, в которых скорости пропорциональны расстоянию до центра симметрии газа; см. также [10]. Дальнейшее уточнение параметров оптимальной системы может быть проведено путем численных расчетов по разностным схемам для уравнений газовой динамики.

Следует отметить, что сходные с нашими в [9] приближенные представления полей газодинамических величин достаточно широко используются при упрощении точных аналитических решений уравнений газовой динамики; см., например, [4].

## § 1. Задача разлета двух плоских поршней

**1.1. Постановка задачи.** Пусть дана (см. рис. 1) система из двух несжимаемых плоских поршней массы  $M_1$  и  $M_3$ , разделенных однородным идеальным газом массы  $M_2$  с внутренней энергией  $E_0$  и показателем адиабаты  $\gamma$ . Не ограничивая общности, будем считать, что поршни движутся в идеальной цилиндри-

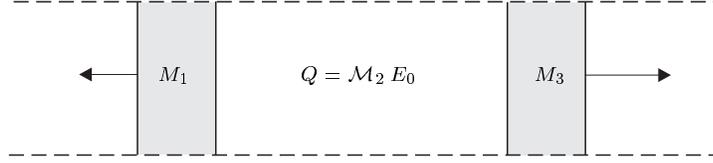


Рис. 1

ческой трубе сечения  $\mathfrak{S} = 1$ . Предполагается, что центр масс системы покоится в лабораторной системе координат. Требуется найти энергию разлетевшихся поршней.

Введем в газе массовую лагранжеву координату  $m$ ,  $m_1 \leq m \leq m_3$ , и выпишем для системы законы сохранения импульса и энергии, считая, что скорость каждого поршня совпадает со скоростью прилегающего газа  $U_3(t) = U(m_3, t)$ ,  $U_1(t) = U(m_1, t)$ :

$$U_3 M_3 + U_1 M_1 + \int_{M_2} U dm = 0, \quad (1.1)$$

$$\frac{1}{2} M_3 U_3^2 + \frac{1}{2} M_1 U_1^2 + \int_{M_2} \frac{1}{2} U^2 dm + \int_{M_2} E_{\text{ВН}} dm = Q, \quad (1.2)$$

где  $Q$  — начальный запас энергии системы, складывающийся из потенциальной энергии газа  $Q_0 = \int_{M_2} E_0 dm$  и начальной кинетической энергии газа и поршней. Рано или поздно, но давление сжатого газа заставит поршни разлетаться, так что асимптотически, в так называемом термодинамическом приближении (см. [4]) полная внутренняя энергия газа

$$\int_{M_2} E_{\text{ВН}}(m, t) dm \approx Q_0 \left[ \frac{x_3(t_0) - x_1(t_0)}{x_3(t) - x_1(t)} \right]^{\gamma-1} \xrightarrow{t \rightarrow \infty} 0; \quad (1.3)$$

здесь  $x_1(t)$  и  $x_3(t)$  — декартовы координаты поршней, и предположено, что во все времена течение оставалось изэнтропическим. Последнее справедливо, если в начальный момент и поршни, и газ находились в покое. Если массой  $M_2$  можно пренебречь по сравнению с  $M_1$  и  $M_3$ , то, поскольку скорости  $U(m, t)$  асимптотически заключены между  $U_1(t)$  и  $U_3(t)$ , в (1.1) и (1.2) третьими членами также можно пренебречь, и мы приходим к простейшей задаче разлета двух материальных точек

$$U_1 M_1 + U_3 M_3 = 0, \quad [M_1 U_1^2 + M_3 U_3^2]_{t=\infty} = 2Q, \quad (1.4)$$

$$\frac{1}{2} M_1 U_1^2|_{t=\infty} = \frac{M_3}{M_1 + M_3} Q, \quad \frac{1}{2} M_3 U_3^2|_{t=\infty} = \frac{M_1}{M_1 + M_3} Q. \quad (1.5)$$

**1.2. Интерполяционная формула при  $\gamma = 3$ .** Примем, что финальная скорость  $U(m, \infty)$  зависит от массовой лагранжевой координаты по линейному закону

$$U(m) = U_3 \frac{m - m_1}{m_3 - m_1} + U_1 \frac{m_3 - m}{m_3 - m_1}. \quad (1.6)$$

Это предположение дает возможность взять интегралы в (1.1) и (1.2):

$$\int_{M_2} U dm = M_2 \frac{U_1 + U_3}{2}, \quad \int_{M_2} \frac{1}{2} U^2 dm = M_2 \frac{U_1^2 + U_1 U_3 + U_3^2}{6}. \quad (1.7)$$

Подставив эти выражения в (1.1) и (1.2), находим

$$U_1 = -\frac{M_3 + M_2/2}{M_1 + M_2/2} U_3, \quad (1.8)$$

$$E_3(\infty) = \frac{1}{2} M_3 U_3^2 \Big|_{t=\infty} = M_3 (M_1 + M_2/2)^2 Q \{ (M_3 + M_2/3)(M_1 + M_2/2)^2 - \\ - M_2 (M_3 + M_2/2)(M_1 + M_2/2)/3 + (M_1 + M_2/3)(M_3 + M_2/2)^2 \}^{-1}. \quad (1.9)$$

Проверим пригодность этой формулы в предельном случае истечения первоначально покоящегося газа в вакуум,  $M_3 = 0$ ,

$$|U_3|^2 = 6E_0 \left[ 1 + \frac{M_1 M_2}{4(M_1 + M_2/2)^2} \right]. \quad (1.10)$$

Точное решение дает (см. [1, § 44] или [4, § 17])

$$U_{\max} = \frac{2}{\gamma - 1} C_0 = \sqrt{\frac{4\gamma}{\gamma - 1} E_0}, \quad C_0 = \sqrt{\gamma(\gamma - 1) E_0}. \quad (1.11)$$

Таким образом, при  $\gamma = 3$  в точном решении  $|U_3|^2 = 6E_0$ . Наша приближенная формула точна в двух крайних случаях:  $M_1 = \infty$  (баллистическая задача Лагранжа) и  $M_1 = 0$  (двустороннее истечение).

В промежутке формула (1.10) слегка уменьшает квадрат скорости, но не более, чем на одну девятую (при  $M_2 = 2M_1$ ); сама скорость истечения лежит между  $(2\sqrt{2}/3)\sqrt{6E_0}$  и  $\sqrt{6E_0}$ .

Кроме того, наше приближенное выражение (1.9) является точным при  $\gamma = 3$  в предельном случае  $M_2 \rightarrow \infty$ . В противоположном чисто кинематическом случае  $M_2 \rightarrow 0$  наше приближенное выражение (1.9) для  $E_3(\infty)$  совпадает с (1.5) и, очевидно, является точным при всех  $\gamma$ . Поэтому можно ожидать, что формула (1.9) как своего рода интерполяционная формула достаточно точна и при промежуточных значениях  $M_2$ .

**1.3. Анализ величины отбора энергии.** Вычислим отношение энергий, отобранных поршнями:

$$\frac{E_3(\infty)}{E_1(\infty)} = \frac{M_3 (M_1 + M_2/2)^2}{M_1 (M_3 + M_2/2)^2}. \quad (1.12)$$

Мы видим, что распределение отобранной поршнями энергии при малых  $M_2$ , обратно пропорциональное их массам (см. (1.5)),

$$E_3 : E_1 \approx M_1 : M_3 \approx |U_3| : |U_1|, \quad (1.13)$$

качественно отличается от распределения при больших  $M_2$ , когда оно пропорционально этим массам,

$$E_3 : E_1 \approx M_1 : M_3, \quad |U_3| \approx |U_1|. \quad (1.14)$$

Заметим, однако, что если первоначально газ и поршни покоились, то скорость разлета ограничена:

$$E_3(\infty) \leq 3M_3E_0, \quad U_3(\infty) \leq \sqrt{6E_0}. \quad (1.15)$$

И если упрощенные формулы (1.5) дают большие значения, это свидетельствует лишь о том, что их применили в ситуации, когда они неприменимы. Как уже отмечалось, наша более точная формула (1.9) к завышенной энергии не приводит.

**1.4. Задача метания пластины.** Рассмотренный выше предельный при  $M_3 \rightarrow \infty$  пример истечения первоначально покоящегося газа в вакуум можно интерпретировать так же, как весьма упрощенную задачу о метании пластины. Достаточно в (1.9) положить  $M_1 = 0$ . Тогда

$$E_3(\infty) = Q \cdot \varphi^{\text{газ}}(\kappa), \quad \kappa = \frac{M_2}{M_3}, \quad (1.16)$$

$$\varphi^{\text{газ}}(\kappa) = \frac{3\kappa}{4 + 5\kappa + \kappa^2}. \quad (1.17)$$

Здесь  $\varphi(\kappa)$  характеризует процент отбора энергии. Нетрудно показать, что оптимум использования энергии

$$\max \varphi^{\text{газ}} = \frac{1}{3} \approx 0,333 \quad \text{при} \quad \kappa_{\text{опт}}^{\text{газ}} = 2. \quad (1.18)$$

Станюковичем и Забабахиным эта задача рассматривалась в более реалистической постановке; см. [5, 6]. К неподвижному поршню  $M_3$  прилегает слой конденсированного ВВ, который подрывается по свободной поверхности. Детонационная волна, пройдя по ВВ со скоростью  $D$ , отражается от поршня ударной волной. Пренебрегая изменением энтропии продуктов взрыва при прохождении этой волны (т.е. в акустическом приближении), можно подсчитать, что

$$\varphi^{\text{ВВ}}(\kappa) = \frac{729}{32\kappa^2} \left( 1 + \frac{16}{27}\kappa - \sqrt{1 + \frac{32}{27}\kappa} \right), \quad (1.18)$$

$$\max \varphi^{\text{ВВ}} = \frac{256}{729} \approx 0,351 \quad \text{при} \quad \varphi_{\text{опт}}^{\text{ВВ}} = \frac{81}{32} \approx 2,53. \quad (1.19)$$

Следующая таблица показывает хорошее согласие нашей приближенной формулы с более точной (1.18):

$\kappa$	1,0	1,2	2,0	2,5	3,0	3,5	4,0	4,5	10	20
ВВ	0,296	0,332	0,347	0,351	0,349	0,345	0,338	0,330	0,225	0,176
газ	0,300	0,315	0,333	0,330	0,322	0,311	0,300	0,289	0,195	0,120

Как было отмечено выше (см. (1.10)), наши формулы уменьшают скорость разлета газа в пустоту. Так как центр масс системы неподвижен, то примерно такой же ошибки следует ожидать и в скорости пластины. Кроме того, более низкие значения  $\varphi^{\text{газ}}$  связаны с различием в постановке задач. Бегущая со

скоростью  $D$  детонационная волна ускоряет продукты взрыва до скорости  $U = D/2$ , и истечение в вакуум в конце заряда происходит со скоростью  $D$  (и со скоростью  $-D/2$  в подрываемой поверхности) вместо значения  $D\sqrt{6}/4$  при объемной детонации. Поэтому и асимптотика отбора энергии  $8\kappa^{-1}$  при больших  $\kappa$  у Станюковича и Забабахина более, чем вдвое превышает асимптотическое значение  $4\kappa^{-1}$  для случая объемной детонации, которое точно воспроизводится в пределе формулой (1.17).

**1.5. Проектирование метания поршня.** Пусть нам требуется построить систему с фиксированной пластиной погонной массы  $M_3$  (приходящейся на единицу площади), приобретающей фиксированную скорость  $U_3$  и, следовательно, фиксированную погонную кинетическую энергию  $E_3$ . Формула (1.9) позволяет строить такие системы, отличающиеся друг от друга количеством газа данной удельной энергии  $E_0$  и погонной массой  $M_1$  “подпорной” пластины. Для этого подставим в нее  $Q = M_2 E_0$  и получим из (1.9) квадратное уравнение относительно  $M_1$ . Разумеется, при этом слишком заметное недоложение газа нельзя будет компенсировать никаким увеличением толщины “подпорной” пластины.

Пусть  $M_2$  — максимальное “количество” газа,  $M_2(\tau) = \tau M_2$  по определению. Тогда в приближении  $M_2 \ll M_1(1)$ ,  $M_3$  для  $M_1(\tau)$  получаем линейное уравнение, из которого

$$M_1(\tau) = \frac{M_3 M_1(1)}{\tau[M_1(1) + M_3] - M_1(1)} \quad (1.20)$$

с условием возможности проектирования

$$\tau > \frac{M_1(1)}{M_1(1) + M_3}. \quad (1.21)$$

Другими словами, минимальное количество газа должно быть не меньше той доли газа исходной системы с  $\tau = 1$ , которая отдает свою энергию пластине  $M_3$  согласно (1.5).

Если в (1.9) будет  $M_2 \ll 6M_1$ , то, приняв  $M_1 + M_2/3 \approx M_1 + M_2/2$ , получим более точное соотношение

$$M_1(\tau) = \frac{[M_3 + 2M_2/3 + M_2^2/(12M_3)]M_1(1) + (1 - \tau)R}{\tau[M_3 + 2M_2/3 + M_2^2/(12M_3)] - (1 - \tau)[M_1(1) + M_2/2]}, \quad (1.22)$$

$$R = \left[ \frac{1 + \tau}{2} M_3 + \frac{7\tau}{12} M_2 \right] M_2 - \left[ \frac{4 - 3\tau}{6} M_2 + \frac{1 + \tau}{12} \frac{M_2^2}{M_3} \right] M_1(1),$$

где условие возможности построения — неотрицательность знаменателя при рассматриваемых значениях  $\tau$ :

$$\tau > \frac{M_1(1) + M_2/2}{M_3 + M_1(1) + 7M_2/6 + M_2^2/(2M_3)}.$$

**1.6. Интерполяционные формулы при  $\gamma \neq 3$ .** Простота и относительная точность приближенной формулы отбора энергии, полученной при  $\gamma = 3$  простейшей интерполяцией, вызывают желание построить аналогичные формулы при других значениях  $\gamma$ . Особый интерес могли бы иметь такие формулы при

специальных значениях  $\gamma = 3, 5/3, 7/5, 9/7, 11/9, \dots$  вида  $\gamma = (2n + 1)/(2n - 1)$ ,  $n = 1, 2, 3, \dots$ . Рассмотрим сперва баллистическую задачу Лагранжа

$$M_1 = \infty, \quad \mathcal{U}_1 = 0, \quad \frac{1}{2}M_1 u_1^2 = 0; \quad 0 \leq m_1 \leq m \leq m_3 = M_2, \quad (1.22)$$

и примем, что линейной по массовой координате является степень

$$[u(m)]^\nu = \frac{m}{M_2} (u_3)^\nu, \quad u(m) = \left(\frac{m}{M_2}\right)^{1/\nu} u_3. \quad (1.23)$$

Выберем

$$\nu = 2 \frac{\gamma - 1}{\gamma + 1} = \frac{1}{n}. \quad (1.24)$$

При таком выборе кинетическая энергия газа равна

$$\int_0^{M_2} \frac{1}{2} [u(m)]^2 dm = \left(2 + \frac{4}{\nu}\right)^{-1} M_2 U_3^2 = \frac{M_2 U_3^2}{(4n + 2)}, \quad (1.25)$$

что при подстановке в (1.2) с учетом (1.3) и (1.22) приводит при  $M_3 = 0$  к правильной скорости истечения газа в пустоту:

$$u_3^2 = (u_{\max})^2 = \frac{4\gamma}{\gamma - 1} E_0 = (4n + 2) E_0. \quad (1.26)$$

Следует заметить, что при истечении газа в пустоту образуется волна разрежения, в которой скорости частиц газа при любом  $\gamma$  линейно зависят от декартовой координаты, так что наши интерполяции по массовой координате на “хвосте” волны разрежения заведомо неточны; они передают лишь “усредненную” картину.

При рассмотрении двустороннего разлета возникает необходимость линейно интерполировать степень знакопеременной функции  $U(m)$ . При нечетных  $n = 2k + 1$ , т.е.  $\gamma = 3, \gamma = 7/5 = 1,4, \gamma = 11/9 \approx 1,2$  и т.д., линейно интерполируются корни нечетной  $n$ -й степени, т.е. применяется аналитическая процедура

$$\int_{m_1}^{m_3} U dm = \frac{M_2}{n+1} \frac{u_3 \sqrt[n]{u_3} - u_1 \sqrt[n]{u_1}}{\sqrt[n]{u_3} - \sqrt[n]{u_1}} = \frac{M_2}{n+1} \sum_{j=0}^n (\sqrt[n]{u_3})^j (\sqrt[n]{u_3})^{n-j}, \quad (1.27)$$

$$\int_{m_1}^{m_3} U^2 dm = \frac{M_2}{n+1} \frac{u_3^2 \sqrt[n]{u_3} - u_1^2 \sqrt[n]{u_1}}{\sqrt[n]{u_3} - \sqrt[n]{u_1}} = \frac{M_2}{n+1} \sum_{j=0}^{2n} (\sqrt[n]{u_3})^j (\sqrt[n]{u_3})^{2n-j}.$$

После подстановки выражений (1.27) в (1.1) и (1.2) из (1.1) для корня

$$h = \sqrt[n]{-\frac{u_1}{u_3}} \quad (1.28)$$

получается уравнение  $n$ -й степени

$$-h^n M_1 + \frac{1}{n+1} (-h^n + \dots + h^2 - h + 1) M_2 + M_3 = 0. \quad (1.29)$$

Решив его и подставив  $U_1 = -h^n U_3$  в (1.2), можно вычислить с учетом (1.3) и (1.29) финальную скорость  $U_3$ :

$$U_3^2 = 2Q(2n+1)\{nM_1h^{2n} + [M_2 + (n+1)(M_1 + M_3)]h^n + nM_3\}^{-1}. \quad (1.30)$$

Напомним, что здесь  $n = 2k + 1$  — нечетное число. Легко проверить, что в пределе  $M_2 \rightarrow \infty$ ,  $Q = M_2 \mathcal{E}_0 \rightarrow \infty$ ,  $U_3 = -U_1 = U_{\max} = \sqrt{(4n+2)\mathcal{E}_0}$ , т.е. скорость истечения в вакуум передается правильно; ср. (1.26). Вместо решения уравнения (1.29) можно искать положительный корень у уравнения степени  $n+1$

$$M_1(h^{n+1} + h^n) + \frac{1}{n+1}(h^{n+1} - 1)M_2 - M_3(h+1) = 0, \quad (1.31)$$

которое отличается от (1.29) лишним корнем  $h = -1$ .

В задаче метания пластины приведенные формулы упрощаются:

$$\varkappa = \frac{(n+1)(h+1)}{h^{n+1} - 1}, \quad \varphi(\varkappa) = \frac{2n+1}{n + (\varkappa + n+1)h^n}, \quad (1.32)$$

где первое соотношение есть уравнение для  $h = h(\varkappa)$  (с лишним корнем  $h = -1$ ), а второе — параметрическое задание функции  $\varphi(\varkappa)$ .

Значительно сложнее объяснить интерполяцию при четных  $n$ . Например, в простейшем случае  $n = 2$ ,  $\gamma = 5/3$  приходится интерполировать нечетную вещественную функцию

$$x^{\langle(1/2)\rangle} = \sqrt{|x|} \operatorname{sgn} x = \lim_{k \rightarrow \infty} x \uparrow \frac{2k+1}{4k+3}$$

с особенностью в нуле и затем применять обратную к ней. Мы приведем только расчетные формулы. В соответствии с (1.28)  $h = \sqrt{-U_1/U_3}$  будет единственным положительным корнем аналога уравнения (1.31):

$$\left(M_1 + \frac{1}{3}M_2\right)h^3 + M_1h^2 - M_3h - \left(M_3 + \frac{1}{3}M_2\right) = 0. \quad (1.33)$$

Аналога же уравнения (1.29), которое было бы квадратным уравнением, здесь нет. При этом доля  $(h/(1+h))M_2$  газа движется влево, остальная часть  $(1/(1+h))M_2$  — вправо. Впрочем, последнее утверждение справедливо для всех  $n$ . Финальная скорость поршня  $M_3$  представима в виде

$$U_3^2 = 10Q \left[ \left(5M_1 + \frac{h}{1+h}M_2\right)h^4 + 5M_3 + \frac{1}{1+h}M_2 \right]^{-1}. \quad (1.34)$$

В задаче метания пластины для  $h = h(\varkappa)$ ,  $h > 1$

$$\varkappa = \frac{3(h+1)}{h^3 - 1}, \quad \varphi(\varkappa) = \frac{5}{2 + (\varkappa + 3)h^2} = \frac{5h^3 - 5}{3h^5 + 5h^3 - 2}. \quad (1.35)$$

Максимум достигается на  $h_{\text{опт}}$  — корне уравнения  $2h^5 - 5h^2 - 3 = 0$ ,  $h_{\text{опт}} = 1,47237$ ,  $\varkappa_{\text{опт}} = 3,38384$ ,  $\max \varphi = 0,31567$ . Нетрудно подсчитать, что при  $\varkappa \rightarrow \infty$  имеем  $h(\varkappa) \rightarrow 1$  и  $\varphi_n(\varkappa) \sim (2n+1)\varkappa^{-1}$ , где  $\gamma = (2n+1)/(2n-1)$ . Для асимптотики  $\varkappa \rightarrow 0$ ,  $h(\varkappa) \rightarrow \infty$ ,  $\varphi_n(\varkappa) \sim (2n+1)\varkappa/(n+1)^2$ .

В заключение приведем таблицу значений  $\varphi_n(\varkappa)$  для  $n = 1, 2, 3$ , т.е. для  $\gamma = 3, 5/3, 7/5$ :

$\kappa$	1,0	1,2	2,0	2,5	3,0	3,5	4,0	4,5	10	20
$\varphi_1$	0,300	0,315	0,333	0,330	0,322	0,311	0,300	0,289	0,195	0,120
$\varphi_2$	0,235	0,254	0,298	0,310	0,315	0,316	0,314	0,310	0,248	0,169
$\varphi_3$	0,193	0,213	0,265	0,283	0,295	0,303	0,307	0,308	0,276	0,205

При  $\gamma \rightarrow 1$ ,  $n \rightarrow \infty$  имеем  $h_{\text{опт}}^{n+1} \sim 1 + \sqrt{2}$ ,  $\kappa_{\text{опт}} \sim (n + 1)\sqrt{2}$ ,  $\max \varphi \sim 1 - \sqrt{2}/2 \approx 0,293$ .

## § 2. Течения Лява с поршнями

**2.1. Автомодельное решение.** Лявом были указаны (см. [3, § 45]) одномерные автомодельные движения политропического газа, описываемые в лагранжевых координатах решениями уравнений изэнтропической газовой динамики с разделяющимися переменными. Эти решения принадлежат к классу одномерных движений Седова [10], в которых скорость частицы пропорциональна расстоянию до центра симметрии газа. Постановка Лява совпадает с постановкой баллистической задачи Лагранжа за единственным исключением: в начальный момент неподвижный изэнтропический газ имеет некоторую переменную плотность  $\rho$  (и давление  $p$ , и скорость звука, и т.п.). Газ Лява (как, впрочем, и газ Седова) можно обрезать, заменяя отсеченные объемы специально подобранными поршнями так, что движение оставшегося объема сохраняется. Мы, следуя [3] и [10], выпишем решения, чтобы проанализировать характер выхода движения поршней на предельный режим. Исходной для нас будет система

$$\frac{\partial u}{\partial t} + \frac{\partial p}{\partial m} = 0, \quad \frac{\partial v}{\partial t} - \frac{\partial u}{\partial m} = 0, \quad (2.1)$$

$$pv^\gamma = A, \quad \mathcal{E} = \frac{1}{\gamma - 1}pv, \quad (2.2)$$

где  $m$  — массовая лагранжева координата,  $V = \rho^{-1}$ ,  $\mathcal{E}$  — удельная энергия,  $A$  — характеристика энтропии. Предполагается, что каждый поршень разгоняется давлением газа, что дает граничные условия

$$M_1 \frac{\partial u}{\partial t} \Big|_{m_1} = -p(m_1, t), \quad M_3 \frac{\partial u}{\partial t} \Big|_{m_3} = p(m_3, t). \quad (2.3)$$

Пусть в решении переменные разделяются:

$$u(m, t) = U(t)u(m), \quad p(m, t) = P(t)p(m), \quad v(m, t) = V(t)v(m), \quad (2.4)$$

$$U(0) = 0, \quad U(\infty) = 1; \quad V(0) = P(0) = 1, \quad (2.5)$$

т.е.  $p(m)$ ,  $v(m)$  и  $\mathcal{E}$  — начальные распределения давления, удельных объема и энергии,  $u(m)$  — финальное распределение скоростей,  $1 - [u(t)]^2$  — искомый относительный недобор энергии поршнем.

Системы для “больших” и “малых” величин имеют по первому интегралу

$$U^2 + V^{-\gamma+1} = 1, \quad \frac{\gamma - 1}{4\gamma}u^2 + \mathcal{E} = E, \quad (2.6)$$

где константа интегрирования  $E = \max \mathcal{E}$ . Первый интеграл порожден законом сохранения энергии, второй — аналог соотношения Бернулли, где для упрощения перешли от энтальпии к энергии. Если перейти к новой лагранжевой координате  $y$  — эйлеровой координате при  $t = 0$ , — то

$$u(y) = \mathcal{U} \frac{y}{L}, \quad \mathcal{E}(y) = E \left[ 1 - \left( \frac{y}{L} \right)^2 \right], \quad (2.7)$$

где  $-L \leq y \leq L$ ,  $\max u = \mathcal{U} = \sqrt{4\gamma E/(\gamma-1)}$  — скорость истечения в пустоту газа с удельной энергией  $E$ , начало отсчета эйлеровой координаты взято в точке, где  $u(m) = 0$ . Соответственно

$$\rho(y) = \left[ 1 - \left( \frac{y}{L} \right)^2 \right]^{1/(\gamma-1)} \rho(0), \quad p(y) = \left[ 1 - \left( \frac{y}{L} \right)^2 \right]^{\gamma/(\gamma-1)} p(0). \quad (2.8)$$

Если  $z(t)$  — эйлерова координата фронта газа,  $z(0) = L$ ,  $\dot{z}(t) = \mathcal{U}v(t)$ , то

$$V = \frac{\dot{z}}{\mathcal{U}}, \quad 1 - V^2 = \left[ \frac{L}{z(t)} \right]^{\gamma-1}. \quad (2.9)$$

Временная же зависимость вычисляется из квадратуры

$$\frac{\mathcal{U}}{L} dt = [1 - V^{-\gamma+1}]^{-1/2} dV, \quad (2.10)$$

откуда  $1 - U^2 \sim (\mathcal{U}t/L)^{-\gamma+1}$ .

**2.2. Проектирование поршней под течение.** Более широкий класс автомодельных изэнтропических течений, в которых энтропия уже не обязана быть постоянной во всем объеме, был указан Седовым [10]. Они характеризуются линейной (ср. (2.7)) зависимостью скорости частиц от эйлерова расстояния до центра симметрии газа:

$$u(m, t) = \frac{\dot{z}(t)}{z(t)} x(m, t), \quad x(m, t) = \frac{z(t)}{L} y(m), \quad (2.11)$$

где  $y(m) = x(m, 0)$ . Пусть дано решение Седова для конечного объема газа, естественно, с давлением, обращаемся в ноль на границе. Выясним, какие массы  $M_1$  и  $M_3$  должны быть у поршней, примыкающих к урезанному объему газа в точках  $m = m_1$ ,  $y(m_1) = y_1$  и  $m = m_3$ ,  $y(m_3) = y_3$ , чтобы граничные соотношения (2.3) для решения выполнялись тождественно. Из первого уравнения (2.1) находим для изменения количества движения

$$M_1 \frac{\partial u}{\partial t} \Big|_{m_1} = - \int^{m_1} \frac{\partial p}{\partial m} dm = \int^{m_1} \frac{\partial u}{\partial t} dm, \quad M_3 \frac{\partial u}{\partial t} \Big|_{m_3} = \int_{m_3} \frac{\partial u}{\partial t} dm.$$

Подставляя сюда  $U(m, t) = \dot{z}(t) y(m)/L$ , получаем

$$y_1 M_1 = \int_{-L}^{y_1} y \rho(y) dy, \quad y_3 M_3 = \int_{y_3}^L y \rho(y) dy, \quad (2.12)$$

т.е. центральный статический момент поршня равен таковому же моменту убранного из-за него объема газа. Рассмотрим теперь цилиндрические ( $\nu = 1$ ) и сферические ( $\nu = 2$ ) течения Седова с симметрией (2.11). Для них уравнения движения газа и поршня имеют вид

$$r^{-\nu} \frac{\partial u}{\partial t} + \frac{\partial p}{\partial m} = 0, \quad r^{-\nu} M \frac{\partial u}{\partial t} \Big|_{m_*} = 2\pi\nu p(m_*, t), \quad (2.13)$$

где учтено, что при движении поршень, сохраняя свои массу, изменяет свою погонную массу  $M_{\text{пор}}$  — массу единицы своей площади.

$$(y_*)^{1-\nu} M = 2\nu\pi \int_{m_*} y^{1-\nu} dm = 2\nu\pi \int_{y_*}^L y \rho(y) dy, \quad y = r(m, 0). \quad (2.14)$$

Таким образом, в цилиндрическом случае масса поршня должна быть равна массе убираемого газа, а в сферическом — должны совпадать минус первые моменты.

В течении Лява нетрудно подсчитать финальную энергию, которую приобретут поршни

$$\frac{1}{2} M_1 u_1^2 = -\frac{y_1}{L^2} \frac{p(y_1)}{\gamma - 1}, \quad \frac{1}{2} M_3 u_3^2 = \frac{y_3}{L^2} \frac{p(y_3)}{\gamma - 1}. \quad (2.15)$$

**2.3. Течения Лява при  $\gamma = 3$ .** Для специальных значений  $\gamma = (2n+1)/(2n-1)$ ,  $n$  целое, правая часть (2.10) интегрируется в элементарных функциях. При  $\gamma = 3$

$$\frac{U}{L} t = \sqrt{V^2 - 1}, \quad V^2 = 1 + \left(\frac{Ut}{L}\right)^2, \quad (2.16)$$

$$1 - U^2 = \frac{L^2}{z^2}, \quad z^2 = L^2 + U^2 t^2. \quad (2.17)$$

Заметно упрощаются и выражения для плотности и давления, поскольку  $1/(\gamma - 1) = 1/2$ ,  $\gamma/(\gamma - 1) = 3/2$ . В элементарных функциях вычисляются интегралы, давшие массу  $\mathcal{M}(m_1, m_3)$ , полную начальную внутреннюю энергию  $E_{\text{ВН}}(m_1, m_3)$  газа, заключенного между двумя поршнями, и его финальную кинетическую энергию  $E_k(m_1, m_3)$ :

$$\mathcal{M}(m_1, m_3) = \rho(0) \left\{ y \sqrt{1 - \left(\frac{y}{L}\right)^2} + \frac{L}{2} \arcsin \frac{y}{L} \right\}_{y_1}^{y_3}, \quad (2.18)$$

$$E_{\text{ВН}}(m_1, m_3) = \frac{\rho(0)}{16} \left\{ y \sqrt{1 - \left(\frac{y}{L}\right)^2} \left[ 11 - 14 \left(\frac{y}{L}\right)^2 \right] - 3L \arcsin \frac{y}{L} \right\}_{y_1}^{y_3}, \quad (2.19)$$

$$E_k(m_1, m_3) = \frac{\rho(0)}{16} \left\{ y \sqrt{1 - \left(\frac{y}{L}\right)^2} \left[ 3 - 6 \left(\frac{y}{L}\right)^2 \right] - 3L \arcsin \frac{y}{L} \right\}_{y_1}^{y_3}. \quad (2.20)$$

Для задач проектирования выписанные формулы (2.18)–(2.20) вместе с кубическими уравнениями для  $1 - (y_i/L)^2$ , вытекающими из (2.15), не слишком удобны даже в классической баллистической задаче, т.е. при  $M_1 = \infty$   $y_1 = 0$  (ср. [3]).

Зато очень полезна ориентировка при помощи (2.17) и (2.9). Кроме того, сравнение профилей (2.7) и (2.8) при разных  $\gamma$  в известной мере оправдывает интерполяции § 1 при  $\gamma \neq 3$ .

Заметим еще, что связи (2.16) и (2.17) аналитические и существуют и при  $t < 0$ . Там они описывают обратный процесс изэнтропического сжатия объема газа до того плотного состояния, которое мы принимали в исходной задаче за начальное для разлета. То же самое относится к изэнтропическим цилиндрическим и сферическим течениям Седова. Только там связь (2.10) заменяется (см. [10]) на

$$\frac{U}{L} dt = \frac{dV}{\sqrt{1 - V^{(1+\nu)(1-\gamma)}}}.$$

Таким образом, временные законы (2.16) и (2.17) описывают изэнтропическое сжатие и последующий разлет объема газа с  $\gamma = 5/3$ , заключенного внутри сферической оболочки.

### § 3. Схождение сферического и цилиндрического поршней

**3.1. Постановка задач.** В сферической полости радиуса  $R$  с абсолютно жесткой стенкой помещен сферический поршень объема  $V_{об}$  из несжимаемой жидкости. Между поршнем и стенкой впущен однородный политропический газ давления  $p_0$  с показателем адиабаты  $\gamma$  (рис. 2). В начальный момент система покоится. Требуется оценить энергию  $E$ , которую отберет поршень на момент схлопывания, если его начальный внешний радиус равен  $r_0$ . Далее, требуется найти, при каком начальном положении  $r_0$  отобранная энергия будет максимальной.

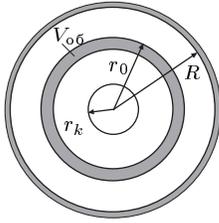


Рис. 2

Аналогично ставится цилиндрическая задача, только вместо объема  $V_{об}$  задается площадь  $S_{об}$  нормального сечения, и идет речь об оценке погонной энергии  $E$ , приходящейся на “отрезок” поршня единичной длины.

Обозначим через  $r_{\Phi}$  внешний радиус поршня в момент схлопывания

$$\frac{4}{3}\pi r_{\Phi}^3 = V_{об} \quad \text{или} \quad \pi r_{\Phi}^2 = S_{об}. \quad (3.1)$$

В силу несжимаемости материала поршня радиус  $r_{\Phi}$  не зависит от начального положения оболочки, а только от ее начального объема  $V_{об}$ . Примем, что давление газа на момент схлопывания выравнено по всему объему, а его кинетическая энергия в этот момент пренебрежимо мала. Эти предположения справедливы, если масса оболочки настолько велика, что она разгоняется достаточно медленно, а плотность ее много больше плотности газа. Добавим, что в сферическом случае область, где скорости газа заметны, примыкает к оболочке и занимает потому весьма малый объем (да и сам наружный край оболочки перед финалом притормаживает). Последнее замечание справедливо в известной мере и для цилиндрического поршня. Однако для плоского аналога задачи уже неразумно пренебрегать кинетической энергией газа, а следует ее учесть с помощью интерполяции из § 1. Так или иначе, здесь мы принимаем, что имеет место баланс

энергий

$$\mathcal{E}_0 = \mathcal{E}_\Phi + E(R, r_0), \quad (3.2)$$

где  $E(R, r_0)$  — искомая кинетическая энергия, отобранная оболочкой на момент схлопывания.

**3.2. Оптимальная сферическая система.** В сферической задаче газ в начальный момент обладает внутренней энергией

$$\mathcal{E}_0 = \frac{4}{3}\pi(R^3 - r_0^3) \frac{p_0}{\gamma - 1}. \quad (3.3)$$

К моменту схлопывания в газе останется, согласно сделанному термодинамическому приближению,

$$\mathcal{E}_\Phi = \frac{4}{3}\pi(R^3 - r_\Phi^3) \frac{p_\Phi}{\gamma - 1} = \mathcal{E}_0 \left( \frac{R^3 - r_0^3}{R^3 - r_\Phi^3} \right)^{\gamma - 1}. \quad (3.4)$$

Тогда из (3.2)

$$E(R, r_0) = \frac{4}{3}\pi(R^3 - r_0^3) \frac{p_0}{\gamma - 1} \left[ 1 - \left( \frac{R^3 - r_0^3}{R^3 - r_\Phi^3} \right)^{\gamma - 1} \right], \quad (3.5)$$

где множитель в квадратных скобках есть доля отобранной поршнем энергии. Ясно, что он  $\sim 1$  при  $r_0 \sim R$ , но тогда мал сам запас энергии  $\mathcal{E}_0$ , пропорциональный  $R^3 - r_0^3$ . При малых же  $r_0$  мал коэффициент отбора. Продифференцировав правую часть (3.5) по параметру  $r_0$  и приравняв производную нулю, находим оптимальный радиус из уравнения

$$R^3 - r_{\text{опт}}^3 = (R^3 - r_\Phi^3) \gamma^{-1/(\gamma - 1)}. \quad (3.6)$$

Следовательно,

$$r_{\text{опт}}^3 = R^3 (1 - \gamma^{-1/(\gamma - 1)}) + r_\Phi^3 \gamma^{-1/(\gamma - 1)}, \quad (3.7)$$

$$E_{\text{max}} = \frac{4}{3}\pi \frac{p_0}{\gamma - 1} (R^3 - r_\Phi^3) \frac{\gamma - 1}{\gamma^{1 + 1/(\gamma - 1)}} = \frac{4}{3}\pi \frac{p_0}{\gamma} (R^3 - r_{\text{опт}}^3). \quad (3.8)$$

Для показателя  $\gamma = 3$  имеем

$$r_{\text{опт}}^3 = 0,42265R^3 + 0,57735r_\Phi^3 \geq (0,75046R)^3, \quad (3.9)$$

$$E_{\text{max}} = \frac{2}{3}\mathcal{E}_0(R, r_{\text{опт}}) = \frac{2}{3\sqrt{3}}\mathcal{E}_0(R, r_\Phi). \quad (3.10)$$

Интересно, что (3.5) показывает также невозможность заметно увеличить отбор энергии оболочкой фиксированного радиуса  $r_0$  лишь за счет увеличения размеров полости и соответствующего увеличения количества сжатого газа. Для  $R \rightarrow \infty$  при фиксированном  $r_0$  из асимптотики коэффициента отбора находим

$$E(\infty, r_0) = \frac{4}{3}\pi(r_0^3 - r_\Phi^3)P_0. \quad (3.11)$$

Чтобы увеличить отбор энергии, можно построить двухкаскадную систему; см. [6, 7]. Тогда разогнанный поршень наружного каскада добавит энергию в газ, сжав его (см. [9]) так, что отбор энергии внутренней оболочкой может заметно увеличиться. Хотя наружный поршень в процессе воздействия будет двигаться в обе стороны: сперва внутрь, затем наружу, поведение внутреннего каскада в этом процессе близко к рассмотренному выше. В частности, формула (3.9) оказывается достаточно точной.

**3.3. Оптимальная цилиндрическая система.** Все следующие формулы аналогичны формулам предыдущего пункта. Газ в начальный момент обладает погонной внутренней энергией

$$\mathcal{E}_0 = \pi(R^2 - r_0^2) \frac{p}{\gamma - 1}. \quad (3.12)$$

Далее, из (3.2) и термодинамического приближения

$$E(R, r_0) = \pi(R^2 - r_0^2) \frac{p_0}{\gamma - 1} \left[ 1 - \left( \frac{R^2 - r_0^2}{R^2 - r_\Phi^2} \right)^{\gamma - 1} \right]. \quad (3.13)$$

Дифференцируя (3.13) по параметру  $r_0^2$ , находим для оптимального радиуса

$$R^2 - r_{\text{опт}}^2 = \frac{R^2 - r_\Phi^2}{\gamma^{1/(\gamma-1)}}. \quad (3.14)$$

Следовательно, для  $r_{\text{опт}}^2$  справедливо (3.7) с заменой кубов на квадраты. Далее,

$$E_{\text{max}} = \pi \frac{p_0}{\gamma^{\gamma/(\gamma-1)}} (R^2 - r_\Phi^2) = \pi \frac{p_0}{\gamma} (R^2 - r_{\text{опт}}^2). \quad (3.15)$$

Для показателя  $\gamma = 3$  вместо (3.9)

$$r_{\text{опт}}^2 = 0,42265R^2 + 0,57735r_\Phi^2 \geq (0,65012R)^2. \quad (3.16)$$

Соотношение (3.10) для  $E_{\text{max}}$  сохраняется. Справедливы и другие сделанные в п. 3.2 замечания.

## § 4. Передача энергии в системе трех поршней

**4.1. Величина передаваемой энергии.** Рассматривается система трех параллельных пластин погонных масс  $M_1$ ,  $M_3$ ,  $M_5$ . Неподвижные в начальный

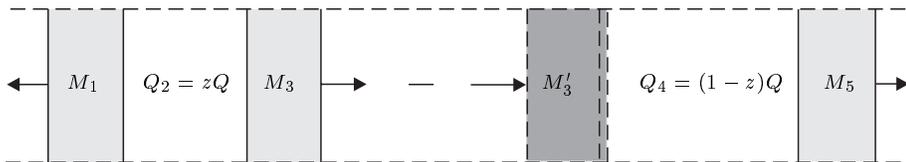


Рис. 3

момент пластины  $M_1$  и  $M_3$  расталкиваются невесомым газом погонной внутренней энергии  $Q_2 = zQ$ . Отобрав всю причитающуюся ей долю энергии, пластина  $M_3$  затем взаимодействует с пластиной  $M_5$  через невесомый газ энергии  $Q_4 = (1 - z)Q$  (см. рис. 3).

Требуется выбрать параметр  $z$  так, чтобы энергия  $E^{(5)} = E_{\text{кин}}^{(5)}(\infty)$  (и, тем самым, предельная при  $t \rightarrow \infty$  скорость  $w_5$ ), приобретаемая пластиной  $M_5$ , имела максимальное значение.

Считая, что пластины  $M_1$  и  $M_3$  уже отобрали всю энергию у газа, из (1.5) получаем

$$u_3 = +\sqrt{\frac{M_1}{M_5} \frac{2zQ}{M_1 + M_3}}. \quad (4.1)$$

Улетев далеко, пластина  $M_3$  образует систему взаимодействия  $M_3Q_4M_5$ . В отличие от исходной  $M_1Q_2M_3$ , система  $M_3Q_4M_5$  имеет ненулевой импульс  $M_3U_3$ , так что ее центр масс движется со скоростью

$$v = \frac{M_3u_3}{M_3 + M_5} = +\sqrt{\frac{M_1M_3}{M_1 + M_3} \frac{2zQ}{(M_3 + M_5)^2}}. \quad (4.2)$$

После взаимодействия пластины  $M_3$  и  $M_5$  начнут расходиться, причем для расчета финальных скоростей  $v_3$  и  $v_5$  в системе центра масс по законам сохранения энергии и импульса

$$\begin{aligned} M_5v_5^2 + M_3v_3^2 &= 2(1 - z)Q + M_5v^2 + M_3(u_3 - v)^2 = \\ &= 2(1 - z)Q + \frac{M_3M_5}{M_3 + M_5} \frac{M_1}{M_3} \frac{2zQ}{M_1 + M_3}; \end{aligned}$$

$$M_5v_5 + M_3v_3 = 0,$$

$$v_5 = +\sqrt{\frac{M_3}{M_5} \frac{1}{M_3 + M_5} \left( 2(1 - z)Q + \frac{M_1M_5}{M_1 + M_3} \frac{2zQ}{M_3 + M_5} \right)}, \quad (4.3)$$

$$w_5 = v_5 + v = \sqrt{B(1 - z) + Az} + \sqrt{Az}, \quad (4.4)$$

$$A = \frac{M_1M_3}{M_1 + M_3} \frac{2Q}{(M_3 + M_5)^2}, \quad B = \frac{M_3}{M_5} \frac{2Q}{M_3 + M_5}. \quad (4.5)$$

**4.2. Оптимизация передачи энергии.** Для определения максимума  $w_5(z)$  продифференцируем правую часть (4.4) по  $z$  и приравняем производную нулю. Получаем

$$\frac{B - A}{A} = \sqrt{\frac{B(1 - z_{\text{опт}}) + Az_{\text{опт}}}{Az_{\text{опт}}}}, \quad (4.6)$$

$$z_{\text{опт}} = \frac{A}{B - A} = \frac{1}{B/A - 1} = \frac{M_1M_5}{M_3(M_1 + M_3 + M_5)}. \quad (4.7)$$

По физическому смыслу задачи точка  $z_{\text{опт}}$  дает экстремум, если  $0 \leq z_{\text{опт}} \leq 1$ . В противном случае экстремум на границе. Правая часть (4.7) неотрицательна,

т.е.  $z_{\text{опт}}$ . Неравенство  $z_{\text{опт}} \leq 1$  эквивалентно  $B/A \geq 2$ , т.е.

$$M_1 + M_3 + M_5 \geq \frac{M_1 M_5}{M_3}. \quad (4.8)$$

Будем считать это неравенство выполненным и, тем самым, обозначение корня через  $z_{\text{опт}}$  оправданным. Тогда, подставив (4.7) в (4.4), имеем

$$\max w_5 = \frac{B}{\sqrt{B-A}}, \quad E_{\text{max}}^{(5)} = \frac{M_1 + M_3}{M_1 + M_3 + M_5} Q. \quad (4.9)$$

Другими словами, эта максимальная энергия совпадает с энергией, которую согласно (1.5) отобрала бы пластина погонной массы  $M_5$  в системе с невесомым газом энергии  $Q = Q_2 + Q_4$  и пластиной погонной массы  $M_1 + M_3$ , т.е. с пластиной, получаемой сложением пластин. При этом окончательная скорость пластины  $M_3$  совпадает со скоростью пластины  $M_1$ :  $w_{3, \text{опт}} = v_{3, \text{опт}} + v_{\text{опт}} = U_{1, \text{опт}}$ , и со скоростью объединенной пластины  $M_1 + M_3$ , редуцированной системе. Таким образом, плоские “двухкаскадные системы” с невесомым газом выигрыша не дают, и при специальных “оптимальных” условиях не дают и проигрыша. Этот вывод справедлив, пока не учитывается масса газа и, следовательно, скорость звука считается бесконечно большой. Тогда скорость поршня в однокаскадной системе не может превзойти скорости истечения газа в вакуум, а в двухкаскадной — уже может; см. [7]. Однако и там существует оптимальный для данных масс  $M_1, M_3, M_5$  расклад разгоняющего газа, и (4.7) дает полезный ориентир. Разумеется, в конкретных системах надо учитывать неполноту отбора энергии в смысле (2.9) и (2.15) или маневрировать массами поршней.

Для полноты рассмотрим случай, когда условие (4.8) не выполнено, и корень  $z_{\text{опт}} > 1$ . Так как  $w_5'(z) > 0$  при малых  $z$ , то экстремум — на границе  $z = 1$ , т.е. выгодно, разогнав всем газом  $Q$  пластину  $M_3$  в системе  $M_1 Q M_3$ , ударить затем ею непосредственно по пластине  $M_5$ . В этом случае

$$w_5 = 2\sqrt{A}, \quad E_{\text{max}}^{(5)} = \frac{4M_1 M_3 M_5}{(M_1 + M_3)(M_3 + M_5)^2} Q. \quad (4.10)$$

Совершенно аналогично можно рассмотреть случай, когда погонная энергия газа  $Q_2$  выше, чем у  $Q_4$ ,  $Q_4 = (1-z)Q$ ,  $Q_2 = kzQ$ ,  $k > 1$ . Приведем окончательные формулы при условии  $0 \leq z_{\text{опт}} < 1$ :

$$z_{\text{опт}} = \frac{1}{B/(kA) - 1}, \quad (4.11)$$

$$E_{\text{max}}^{(5)} = \frac{M_3(M_1 + M_3)}{M_3(M_1 + M_3 + M_5) - (k-1)M_1 M_5} Q. \quad (4.12)$$

Если корень  $z_{\text{опт}} < 0$  или  $z_{\text{опт}} \geq 1$ , то физический максимум достигается при  $z = 1$ :

$$E_{\text{max}}^{(5)} = \frac{4M_1 M_3 M_5}{(M_1 + M_3)(M_3 + M_5)} kQ. \quad (4.13)$$

## СПИСОК ЛИТЕРАТУРЫ

1. *Courant R., Friedrichs K.O.* Supersonic Flow and Shock Waves. — N.Y.: Interscience, 1948; [Рус. пер.: Курант Р., Фридрихс К.О. Сверхзвуковое течение и ударные волны. — М.: ИЛ, 1950].
2. *Poisson S.D.* Formules relatives au Mouvement du Boulet . . . extraites des Manuscrite de Lagrange // Journ. de l'Ecole Polytechnique. 1832. Cahier 21.
3. *Love A.E.H., Pidduck F.B.* Lagrange's Ballistic Problem // Philosoph. Transact. of the Royal Soc. of London. Ser. A. 1922. V. 222. P. 167–226.
4. *Станюкович К.П.* Неустановившиеся движения сплошной среда. — М.: Гостехиздат, 1955. — 2-изд. — М.: Наука, 1971.
5. *Баум Ф.А., Станюкович К.П., Шехтер Б.И.* Физика взрыва. — М.: Физматгиз, 1959.
6. *Забабахин Е.И., Забабахин И.Е.* Неограниченная кумуляция энергии. — М.: Наука, 1988.
7. *Забабахин Е.И.* Явления неограниченной кумуляции // Механика в СССР за 50 лет. Т. 2. — М.: Наука, 1970. — С. 313–342.
8. *Брушлинский К.В., Каждан Я.М.* Об автомодельных решениях некоторых задач газовой динамики // Успехи матем. наук. 1963. Т. 18, вып. 2. С. 3–23.
9. *Забродин А.В., Ченцов Н.Н.* Некоторые приближенные газодинамические модели в задачах передачи и кумуляции энергии. — Забабахинские научные чтения, 1987.
10. *Седов Л.И.* Об интегрировании уравнений одномерного движения газа // Докл. АН СССР. 1953. Т. 90, № 5. С. 735.; Методы подобия и размерности в механике. — Изд. 3-е. — М.: Гостехиздат, 1954.

## ПРИБЛИЖЕННЫЕ РЕШЕНИЯ В БАЛЛИСТИЧЕСКОЙ ЗАДАЧЕ ЛАГРАНЖА И ЕЕ ОБОБЩЕНИЯХ \*)

1. При проектировании газодинамических течений с нужными свойствами определенный интерес представляют компактные приближенные формулы, которые позволяли бы свободно ориентироваться в решении. Тогда расчеты по разностным схемам использовались бы только для уточнения параметров оптимальной системы.

Например, в проблеме кумуляции энергии интересна следующая задача.

В сферической полости радиуса  $R$  с абсолютно жесткой стенкой помещен сферический поршень объема  $V_{об}$  из несжимаемой жидкости. Между поршнем и стенкой впушен однородный политропический газ давления  $P_0$  с показателем адиабаты  $\gamma$ . В начальный момент система покоится. Требуется оценить энергию  $E = E(R, r_0)$ , которую отберет поршень на момент схлопывания, если его начальный внешний радиус равен  $r_0$ , а также найти, при каком начальном положении отобранная энергия будет максимальной.

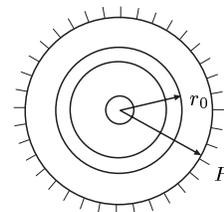


Рис. 1

\*) Конструирование алгоритмов и решение задач математической физики. — М.: ИПМ, 1989. — С. 187–192 (совм. с А.В. Забродиным).

Если принять, что в момент схлопывания давление в газе выровнено, а его кинетическая энергия пренебрежимо мала, можно получить (см. [1]) простую формулу

$$E = (R, r_0) = \mathcal{E}_{\text{нач}}(R, r_0) \left[ \left( \frac{R^3 - r_0^3}{R^3 - r_\varphi^3} \right)^{\gamma-1} \right], \quad (1)$$

$$\mathcal{E}_{\text{нач}}(R, r_0) = \frac{4}{3} \pi (R^3 - r_0^3) \frac{P_0}{\gamma - 1}, \quad (2)$$

где  $r_\varphi$  — финальный радиус оболочки,  $4\pi r_\varphi^3/3 = V_{\text{об}}$ . Отсюда

$$r_{\text{опт}}^3 = R^3 [1 - \gamma^{-1/(\gamma-1)}] + r_\varphi^3 \gamma^{-1/(\gamma-1)}, \quad (3)$$

$$E_{\text{max}} = \frac{4}{3} \pi \frac{P_0}{\gamma} (R^3 - r_{\text{опт}}^3) = \mathcal{E}_{\text{нач}}(R, r_\varphi) \frac{\gamma - 1}{\gamma^{1+1/(\gamma-1)}}. \quad (4)$$

**2.** Рассмотренная задача является сферическим обобщением классической баллистической задачи Лагранжа; см. [2, 3]. Именно, бесконечная цилиндрическая труба перегородена казенником — поршнем  $M_\varepsilon$  бесконечной массы. Правее помещено ядро — поршень  $M_\text{я}$ , могущий свободно скользить в трубе. Камера между казенником и ядром заполнена массой  $\mathcal{M}$  однородного газа удельной энергии  $\varepsilon_0$  с показателем адиабаты  $\gamma$ .

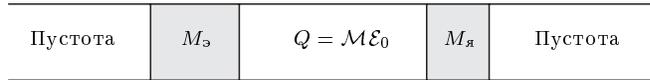


Рис. 2

В начальный момент система покоится. Требуется найти финальную скорость  $u_\text{я}(\infty)$  ядра. Если  $\mathcal{M} \ll M_\text{я}$ , то действует школьная формула  $u_\text{я}(\infty) = \sqrt{2Q/M_\text{я}}$ . В общем случае при фиксированном заряде  $Q = \mathcal{M}\varepsilon_0$  зависимость  $u_\text{я}(\infty)$  от  $M_\text{я}$  имеет (см. [3]) сложный кусочно-аналитический вид. Еще сложнее двусторонняя задача разлета двух поршней; см. [4].

Прибегнем к интерполяции. Введем в газе  $\mathcal{M}$  массовую лагранжеву координату  $t$ , где точка  $t = 0$  отвечает центру масс системы, и предположим, что

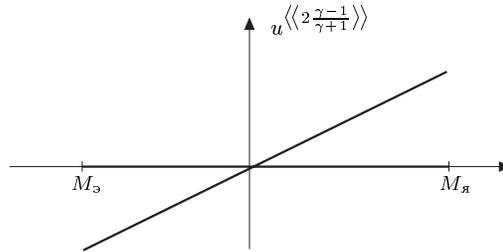


Рис. 3

финальная скорость частицы газа зависит от  $t$  линейно; см. рис. 3. Подставив

эту зависимость в законы сохранения

$$M_{\circ}u_{\circ} + \int_{\mathcal{M}} u \, dm + M_{\text{я}}u_{\text{я}} = 0, \quad (5)$$

$$M_{\circ}u_{\circ}^2 + \int_{\mathcal{M}} u^2 \, dm + M_{\text{я}}u_{\text{я}}^2 = 2Q, \quad (6)$$

можно получить явное выражение для  $u_{\text{я}}$ . Имеем

$$u_{\circ} = -\frac{\widetilde{M}_{\text{я}}}{M_{\circ}} u_{\text{я}}, \quad \text{где } \widetilde{M} = M + \frac{1}{2}\mathcal{M},$$

$$M_{\text{я}}u_{\text{я}}^2 = \frac{2M_{\text{я}}\widetilde{M}_{\circ}^2Q}{(M_{\text{я}} + \mathcal{M}/3)\widetilde{M}_{\circ}^2 - \mathcal{M}\widetilde{M}_{\text{я}}\widetilde{M}_{\circ}/3 + (M_{\circ} + \mathcal{M}/3)\widetilde{M}_{\text{я}}^2}. \quad (7)$$

При  $M_{\text{я}} = 0$ ,  $M_{\circ} = \infty$ ,  $Q = M\mathcal{E}_0$  получаем  $M_{\text{я}}^2 = 6\mathcal{E}_0$ . Как известно (см. [4, 5]), квадрат скорости истечения газа в вакуум  $u_{\text{макс}}^2 = (4\gamma/\gamma - 1)\varepsilon_0$ . Таким образом, наша интерполяция скорости дала точную величину скорости истечения газа в вакуум для случая  $\gamma = 3$ . Аналогично то же значение  $u_{\text{я}}^2 = 6\varepsilon_0$  получается при  $M_{\text{я}} = 0$ ,  $M_{\circ} = 0$ , т.е. при двустороннем истечении. Неплохие результаты получаются и при промежуточных значениях масс.

Обратимся теперь к известной задаче метания пластины, т.е. задаче при  $M_{\circ} = 0$ , рассмотренной Станюковичем [4] и Забабахиным [6]. Только мы рассматриваем объемную детонацию, а у них ВВ подрывалось по свободной поверхности. Отбор энергии  $J(\varkappa)$  в зависимости от отношения масс  $\varkappa = \mathcal{M}/M_{\text{я}}$  виден из таблицы, подсчитанной по формулам из [6] и [1]:

$\varkappa$	1,0	2,0	2,5	3,0	4,0	10,0
[6]	0,296	0,347	0,351	0,349	0,338	0,225
[1]	0,300	0,333	0,330	0,322	0,300	0,195

Заметим, что асимптотики при  $\varkappa \rightarrow \infty$  в [6] и [1] разнятся из-за различия в постановке задач; см. [1, § 1]. Чтобы получить сходные результаты при  $\gamma \neq 3$  надо считать линейной в (5) и (6) величину

$$u^{\langle(2(\gamma-1)/(\gamma+1))\rangle} = |u|^{2(\gamma-1)/(\gamma+1)} \text{sgn } u. \quad (8)$$

**3.** Временные зависимости, описывающие выход  $u_{\text{я}}(t)$  на асимптотику, удобно оценить из автомодельного решения Седова. В нем, в отличие от предыдущего, полагают начальные плотность и удельную энергию переменными:

$$\varepsilon_0(y) = \mathcal{E} \left[ 1 - \left( \frac{y}{L} \right)^2 \right], \quad y \leq L, \quad (9)$$

$$u(y, t) = \mathcal{U} \frac{y}{L} U(t), \quad U(0) = 0, \quad U(\infty) = 1, \quad (10)$$

где лагранжева координата  $y$  частицы есть ее эйлерова координата при  $t = 0$ , а начальная плотность равна

$$\rho_0(y) = \left[ 1 - \left( \frac{y}{L} \right)^2 \right]^{1/(\gamma-1)} \rho_0(0); \quad (11)$$

см. график плотности для случая  $\gamma = 3$  на рис. 4. Оказывается, если заменить заштрихованную часть газа поршнем такой масса  $M_\varepsilon$ , чтобы центр масс системы сохранился в нуле, движение газа остается прежним; см. (10). Аналогично можно подобрать начальное положение ядра  $M_\gamma$ . В классической баллистической задаче Лагранжа последнее заметил еще Ляв в 1922 г.; см. [3, § 45].

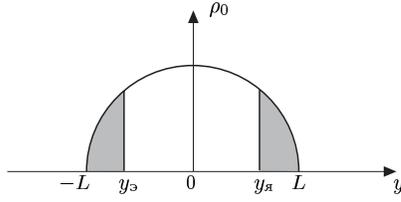


Рис. 4

4. Покажем, что при решении баллистической задачи Лагранжа с  $\gamma \approx 1$  традиционными разностными методами следует пользоваться сеткой, адаптированной к течению. Естественно, постановка задачи подсказывает использовать консервативную разностную схему в лагранжевых координатах. Возьмем вполне консервативную разностную схему, выбрав шаг  $h$  по массовой координате постоянным. Во всех рекомендованных в [7] схемах кинетическая энергия интервала  $\{kh \leq m \leq (k+1)h\}$  есть

$$\frac{1}{2} \left[ \frac{1}{2} u_k^2 + \frac{1}{2} u_{k+1}^2 \right] h. \quad (12)$$

Суммируя, получаем для финальной кинетической энергии

$$h \sum_{k=1}^{n-1} \frac{1}{2} u_k^2 + \frac{1}{4} h u_n^2 + \frac{1}{2} M_\gamma u_n^2 = M \varepsilon_0 = n h \varepsilon_0,$$

которая, по консервативности схемы, равна начальной внутренней энергии газа. Отсюда

$$u_n^2 \leq 4n \varepsilon_0. \quad (13)$$

С другой стороны, при  $M_\gamma = 0$  ответом должна быть скорость истечения газа в вакуум  $u_{\max}^2 = (4\gamma/\gamma - 1)\varepsilon_0$ , т.е. только при числе счетных интервалов  $n > \gamma/(\gamma - 1)$  можно надеяться на малую погрешность расчета. При  $\gamma \approx 1$  это неравенство является объективным требованием, так что при  $\gamma = 1,01$  надо иметь  $n > 101$ .

Таким образом, мы построили контрпример к утверждению [7], что принципы однородности и полной консервативности позволяют “строить схемы, правильно передающие быстро изменяющиеся решения даже на грубых сетках, когда фактически теряется аппроксимация”; см. [7, с. 6]. Как раз при  $\gamma \approx 1$  при малых  $n$  схемное выражение (12) ввиду (8) не аппроксимирует интеграл кинетической энергии газа, и решение получается с большой погрешностью.

Когда, решая уравнения газовой динамики, ищут распределения трех мер — скалярных мер массы и энергии и векторной меры импульса, т.е. суммарных значений указанных величин для различных объемов газа, тогда точность решения естественно оценивать в интегральной метрике  $L_1$ . Но можно искать скорость газа в точке — производную Радона–Никодима от импульс-меры по массе, — и оценивать погрешность построенного поля скоростей  $u(m)$  в метрике  $C$  (или  $L_\infty$ ). При очень легком ядре  $M_\gamma \ll M$  его финальная энергия близка к нулю, и расчет по полностью консервативной схеме этот факт воспроизведет, хотя без

должной адаптации относительная ошибка в финальной энергии и абсолютная ошибка в финальной скорости будут весьма большими.

Интересно выяснить, что стоит за практикой вычислительной газовой динамики, когда выделяют наиболее сильные разрывы и “мажут” разрывы малой интенсивности. Оказывается, что это соответствует измерению погрешности решения в  $D$ -метрике Колмогорова, обобщающей метрику  $C$  на пространство функций с разрывами “первого рода”,

$$\rho_{A,B}(f_1, f_2) = \inf_{Y(\cdot)} \max \left\{ A \sup_x |f_1(x) - f_2(Y(x))|, B \sup_x |x - Y(x)| \right\}, \quad (14)$$

где  $Y(\cdot)$  — взаимно однозначное отображение области задания функций. Пусть разностное решение интерполируется с решетки счетных точек на всю область течения до функции  $f_2(x)$ ; здесь  $x = (x, y, z, t)$ . Затем по (14) оценивают разность истинного решения  $f_1(x)$  и “чуть сдвинутого” оптимальным образом расчета  $f_2(Y(x))$ . Как нам указал после доклада С.К. Годунов, аналогичную метрику предложила использовать М. Шатцман.

#### СПИСОК ЛИТЕРАТУРЫ

1. *Забродин А.В., Ченцов Н.Н.* Препринт // Ин-т прикл. матем. им. М.В.Келдыша АН СССР, 1988, № 58.
2. *Poisson S.D.* // У.Ес. Polytechn. 1832. Cahier 21.
3. *Love A.E.H., Pidduck F.B.* // Trans. Roy. Soc. London, ser. A. 1922. V. 222. P. 167–226.
4. *Станюкович К.П.* Неустановившиеся движения сплошной среды. — М., 1955.
5. *Courant R., Friedrichs K.O.* Supersonic flow and shock waves. — N.Y., 1948.
6. *Забабихин Е.И., Забабихин И.Е.* Явления неограниченной кумуляции. — М., 1988.
7. *Самарский А.А., Попов Ю.П.* Разностные схемы газовой динамики. — М., 1975.

## АППРОКСИМАЦИЯ РЕШЕНИЯ БАЛЛИСТИЧЕСКОЙ ЗАДАЧИ ЛАГРАНЖА \*)

Рассматривается классическая задача о движении без трения в бесконечной цилиндрической трубе сечением  $\sigma$  двух плоских несжимаемых поршней массой  $M_n$  и  $M_p$ , разделенных объемом сильно сжатого идеального газа массой  $M$  с показателем изэнтропы  $\gamma$ . Предполагается, что в начальный момент  $t = 0$  газ однороден и так же, как и поршни, покоится; его удельная внутренняя энергия равна  $\varepsilon_0$ , полная энергия  $Q = M\varepsilon_0$ , плотность  $\rho_0 = Ml^{-1}$ , где  $l = y_p - y_n > 0$ ;  $y_n = x_n(0)$  и  $y_p = x_p(0)$  — начальные декартовы координаты поршней вдоль трубы. Когда  $M_n = \infty$  и, следовательно, скорость  $u_n = 0$ ,  $M_n u_n^2 = 0$ , рассматриваемая задача обращается в задачу Лагранжа [1–3]; когда  $M_n = 0$  — в

\*) ВАНТ. Сер. Матем. моделирование физ. процессов. 1990. В. 3. С. 15–17 (совм. с А.В. Забродиним).

задачу о метании пластины [4, 5]. Аналитическое решение всех этих задач далеко не всегда удается построить явно, и оно имеет сложный вид. В частности, аналитическое решение зависит от того, сколько отражений испытают фронты волн разрежения, возникающих при трогании поршней с места [2, 3]. В то же время простота постановки баллистической задачи делает ее удобной моделью движения при проектировании плоских и даже сферических течений с нужными свойствами (см., например, [4–7]). Поэтому представляют определенный интерес компактные приближенные формулы, позволяющие свободно ориентироваться в решении [4–6]. Дальнейшее уточнение параметров оптимальной системы может быть в принципе проведено путем численных расчетов по какой-либо экономичной разностной схеме для системы уравнений газовой динамики.

В основу грубого приближенного решения обычно кладется то или иное упрощающее предположение. Первое такое предположение было высказано еще в 1793 г. Лагранжем [1], предложившим считать плотность газа постоянной по пространству и, как следствие, скорость — изменяющейся линейно по эйлеровой координате. Как следует из рассмотрения автомодельного течения Пиддака–Кента (указанного впервые, впрочем, Лявом в [2, § 45]), приближение Лагранжа применимо при массах поршней, заметно больших массы газа. Однако ниже рассматривается главным образом случай легких поршней, в том числе задача метания пластины.

Введем в газе массой  $\mathcal{M}$  кроме декартовой массовую лагранжеву координату  $m$  и рассмотрим для изэнтропического [2, 3] разлета поршней финальные скорости  $u(m)$ . Из законов сохранения импульса и энергии следует

$$M_{\text{л}}u_{\text{л}} + M_{\text{п}}u_{\text{п}} + \int u \, dm = 0, \quad (1)$$

$$M_{\text{л}}u_{\text{л}}^2 + M_{\text{п}}u_{\text{п}}^2 + \int u^2 \, dm = 2Q. \quad (2)$$

В предельном случае  $\mathcal{M} \ll \min(M_{\text{л}}, M_{\text{п}})$  интегралами в (1) и (2) можно пренебречь, вследствие чего получаем классические формулы разлета двух масс

$$\frac{u_{\text{л}}}{u_{\text{п}}} = \frac{M_{\text{п}}}{M_{\text{л}}}, \quad \frac{1}{2}M_{\text{л}}u_{\text{л}}^2 = \frac{M_{\text{п}}Q}{M_{\text{л}} + M_{\text{п}}}, \quad \frac{1}{2}M_{\text{п}}u_{\text{п}}^2 = \frac{M_{\text{л}}Q}{M_{\text{л}} + M_{\text{п}}}. \quad (3)$$

В общем случае поместим начало отсчета  $m = 0$  в центр масс системы, т.е. так, чтобы  $u(m) \leq 0$  при  $m_{\text{л}} \leq m \leq 0$  и  $u(m) \geq 0$  при  $0 \leq m \leq m_{\text{п}}$ ,  $m_{\text{п}} - m_{\text{л}} = \mathcal{M}$ . Положим  $w(m) = |u(m)|^{\beta} \operatorname{sgn} u(m)$ , где  $\beta = 2(\gamma - I)(\gamma + I)^{-1}$ ;  $\nu = \beta^{-1}$ ;  $\gamma = (2\nu + I)(2\nu - I)^{-1}$ , и примем интерполяционную гипотезу, что эта “степень” скорости линейно зависит от координаты  $m$ ,  $w(m) = \lambda m$ ,  $\lambda = w_{\text{л}} : m_{\text{л}} = w_{\text{п}} : m_{\text{п}}$ . Тогда

$$-m_{\text{л}} : m_{\text{п}} = (-u_{\text{л}})^{\beta} : u_{\text{п}} = \mu, \quad m_{\text{л}} = -\mu\mathcal{M}(1 - \mu)^{-1}, \quad m_{\text{п}} = \mathcal{M}(1 + \mu)^{-1}, \quad (4)$$

$$\int u \, dm = (1 + \nu)^{-1}[u_{\text{п}}m_{\text{п}} - u_{\text{л}}m_{\text{л}}], \quad \int u^2 \, dm = (1 + 2\nu)^{-1}[u_{\text{п}}^2m_{\text{п}} - u_{\text{л}}^2m_{\text{л}}^2]. \quad (5)$$

Подставляя выражения (5) в (1) и (2), получаем с учетом (4) две формулы

связи  $u_{\text{л}}$  и  $u_{\text{п}}$ :

$$u_{\text{л}}[M_{\text{л}} - (1 + \nu)^{-1}m_{\text{л}}] + u_{\text{п}}[M_{\text{п}} + (1 + \nu)^{-1}m_{\text{п}}] = 0, \quad (6)$$

$$u_{\text{л}}^2[M_{\text{л}} - (1 + 2\nu)^{-1}m_{\text{л}}] + u_{\text{п}}^2[M_{\text{п}} + (1 + 2\nu)^{-1}m_{\text{п}}] = 2Q. \quad (7)$$

Из (6) вытекает “алгебраическое” уравнение для отношения  $\mu$  с единственным допустимым корнем

$$M_{\text{л}}\mu^{\nu} + (1 + \nu)^{-1}\mathcal{M}(\mu^{\nu+1} - 1)(\mu + 1)^{-1} - M_{\text{п}} = 0. \quad (8)$$

Решив его, можно в явной форме выразить из (7) и (4) как обе финальные скорости, так и финальные энергии поршней. Нетрудно проверить, что в предельных случаях  $\mathcal{M} \gg \max(M_{\text{л}}, M_{\text{п}})$  и  $M_{\text{л}} = \infty$ ,  $\mathcal{M} \gg M_{\text{п}}$  финальная скорость разгоняемого поршня достигает теоретического значения скорости нестационарного истечения газа в вакуум

$$u^2 = 4\gamma(\gamma - 1)^{-1}\varepsilon_0 = 2(1 + 2\nu)\varepsilon_0. \quad (9)$$

В случае же  $\mathcal{M} \ll \min(M_{\text{л}}, M_{\text{п}})$  получаются асимптотически точные формулы (3). Поэтому можно ожидать, что описанный интерполяционный алгоритм достаточно точен и при промежуточных соотношениях  $\mathcal{M}$ ,  $M_{\text{л}}$ ,  $M_{\text{п}}$ .

Уравнение для  $\mu$  превращается в целое алгебраическое уравнение степени  $2k - 1$  при  $\nu = 2k - 1$ , т.е. при  $\gamma = 3, 7/5, 11/9, \dots$ . При  $\nu = 1, \gamma = 3$  оно сводится к линейному уравнению и совпадает с классическим приближением Лагранжа

$$-u_{\text{л}} : u_{\text{п}} = M_{\text{п}}^* : M_{\text{л}}^*, \quad M^* = M + \frac{1}{2}\mathcal{M},$$

$$\frac{1}{2}M_{\text{п}}u_{\text{п}}^2 = \frac{M_{\text{п}}(M_{\text{л}}^*)^2Q}{(M_{\text{л}}^*)^2(M_{\text{п}} + \mathcal{M}/3) + (M_{\text{п}}^*)^2(M_{\text{л}} + \mathcal{M}/3) - M_{\text{л}}^*M_{\text{п}}^*\mathcal{M}/6}.$$

Эти формулы в случае  $\gamma = 3$  могут быть применены для проектирования метания пластины при объемном выделении энергии в газе [6, 7]. Положив  $M_{\text{л}} = 0$ ,  $M_{\text{л}}^* = \mathcal{M}/2$ , получаем

$$E_{\text{п}} = \frac{1}{2}M_{\text{п}}u_{\text{п}}^2 = Q\varphi^{\text{газ}}(\varkappa), \quad \varphi^{\text{газ}}(\varkappa) = \frac{3\varkappa}{4 + 5\varkappa + \varkappa^2}, \quad \varkappa = \frac{\mathcal{M}}{M_{\text{п}}}. \quad (10)$$

Здесь  $\varphi(\varkappa)$  характеризует процент отбора энергии. Более интересная постановка задачи метания с бегущей детонационной волной в  $BB$  рассматривалась в работах [4, 5]. Для величины  $\varphi^{BB}$  отбора энергии в них получено

$$\varphi^{BB}(\varkappa) = \frac{729}{32\varkappa^3} \left\{ 1 + \frac{16}{27}\varkappa - \sqrt{1 + \frac{32}{27}\varkappa} \right\}^2.$$

Следующая таблица показывает хорошее согласие обоих подходов при умеренных количествах  $BB$ :

Отбор энергии  $\varphi(\varkappa)$

$\varkappa$	1,0	2,0	2,5	3,0	3,5	4,0	10,0
$\varphi^{BB}(\varkappa)$	0,296	0,347	0,351	0,349	0,345	0,338	0,225
$\varphi^{\text{газ}}(\varkappa)$	0,300	0,333	0,330	0,322	0,311	0,300	0,195

При больших значениях  $\varkappa$  согласия ожидать нельзя из-за различия в постановке задач. Бегущая со скоростью  $D$  детонационная волна ускоряет продукты взрыва до скорости  $u = D/2$ , и истечение в вакуум в конце заряда происходило бы со скоростью  $D$  вместо скорости  $D\sqrt{6}/4$  при объемной детонации, положенной в основу настоящего рассмотрения. Так как легкий поршень полетит почти со скоростью продуктов взрыва, то отборы энергии будут отличаться почти в три раза. Наоборот, при малых  $\varkappa$  асимптотики  $\varphi^{\text{газ}}(\varkappa) \approx 0,75\varkappa$  и  $\varphi^{BB}(\varkappa) \approx 0,7\varkappa$  близки друг к другу. Аналогичные расчеты отбора энергии при других показателях адиабаты  $\gamma$  продуктов взрыва приведены в [6].

В случае произвольного  $\gamma$  финальная скорость явно выражается через корень  $\mu$  уравнения (8)

$$u_{\text{п}}^2 = 2Q [\mu^{2\nu} M_{\text{л}} + (1 + 2\nu)^{-1} \mathcal{M}(\mu^{2\nu+1} - 1)(\mu + 1)^{-1} + M_{\text{п}}]^{-1}.$$

Другой путь приближенного решения баллистической задачи Лагранжа состоит в построении автомодельных течений Лява–Пиддака–Кента (см. [2, § 45] и [8]). Газ Лява в начальный момент неподвижен и изэнтропичен,  $p\rho^{-\gamma} = A = \text{const}$ , но имеет некоторую переменную плотность  $\rho(y)$  и удельную энергию  $\varepsilon_0(y)$

$$\rho(y) = [1 - y^2 L^{-2}]^{1/(\gamma-1)} \rho(0), \quad \varepsilon_0(y) = [1 - y^2 L^{-2}] \mathcal{E}, \quad (11)$$

где  $y$  — эйлерова координата при  $t = 0$  и ее удобно выбрать в газе в качестве лагранжевой. Финальная скорость газа изменяется по  $y$  линейно:

$$u(y) = \mathcal{U}yL^{-1}, \quad \mathcal{U}^2 = 4\gamma(\gamma - 1)\mathcal{E}, \quad -L \leq y \leq L. \quad (12)$$

Временные зависимости  $x(y, t) = yV(t)$ ;  $u(y, t) = U(t)u(y)$ ;  $\rho(y, t) = [V(t)]^{-1}\rho(y)$ , где  $U^2 + V^{1-\gamma} = 1$ , находятся интегрированием уравнения

$$\mathcal{U}L^{-1}dt = [1 - V^{1-\gamma}]^{-1/2}dV, \quad V(0) = 1. \quad (13)$$

Особенно простой вид имеет квадратура (13) при  $\gamma = 3$ :

$$[V(t)]^2 = 1 + \mathcal{U}^2 t^2 L^{-2}, \quad [U(t)]^2 = \mathcal{U}^2 t^2 [V(t)]^{-2}. \quad (14)$$

Течение массы исходного газа с  $y_{\text{л}} \leq y \leq y_{\text{п}}$ , ограниченной поршнями массой  $M_{\text{л}}$  и  $M_{\text{п}}$ , описывается тем же автомодельным решением, если центральный статический момент каждого поршня равен такому же моменту отсеченного им и затем убранного объема газа [2, 6]:

$$y_{\text{л}}M_{\text{л}} = \int_{-L}^{y_{\text{л}}} y\rho(y) dy, \quad y_{\text{п}}M_{\text{п}} = \int_{y_{\text{п}}}^L y\rho(y) dy. \quad (15)$$

Финальная энергия, которую отбирают поршни,

$$\frac{1}{2}M_{\text{л}}u_{\text{л}}^2 = -y_{\text{л}}L^{-2}(\gamma - 1)^{-1}p(y_{\text{л}}), \quad \frac{1}{2}M_{\text{п}}u_{\text{п}}^2 = y_{\text{п}}L^{-2}(\gamma - 1)^{-1}p(y_{\text{п}}).$$

Масса  $M$  газа и его начальная энергия  $Q$  вычисляются по параметрам  $L, \mathcal{E}, A$  интегрированием  $\rho(y)$  и  $\varepsilon_0(y)\rho(y)$  по области  $y_{\text{л}} \leq y \leq y_{\text{п}}$ . Обратный же подбор  $L$

и  $\mathcal{E}$  (или  $\mathcal{U}$ ) по  $M_d, M_n, M$  и  $Q$  сводится к решению трансцендентных уравнений [2, 6]. Тем не менее рассмотрение автомодельного течения дает представление о характере выхода движения на асимптотический режим, а связи (11) и (12) в некоторой мере оправдывают принятую выше интерполяционную гипотезу.

Сопоставляя (15) и (11), видим, что при  $M \ll \min\{M_d, M_n\}$  предположение Лагранжа о постоянстве плотности по пространству практически выполняется. Однако при  $M \gg \min\{M_d, M_n\}$  картина течения у легкого поршня совершенно иная. Если поршни все же достаточно массивны, то на финальной стадии течения может быть близко к указанному автомодельному. По прикидкам Пиддака и Кента (см. [8] и указанную там литературу), для этого нужно, чтобы фронты волн разрежения, возникающих, когда поршни трогаются с начальных мест, испытали несколько отражений от них. Однако уже при  $M_d = \infty$

$$M \geq \frac{2\gamma}{\gamma+1} [2^{(\gamma+1)/2(\gamma-1)} - 1] M_n$$

(если  $\gamma = 3$ , то при  $M \geq 3M_n/2$ ) фронт отразится от казенника только один раз [3]. Классический расчет течения в этой ситуации, данный в [2], весьма сложен (см. также [3] и последующее упрощение методики расчета в [9]).

Использование простых альтернативных приближенных моделей в баллистической задаче вызывается также и тем, что численное ее решение с помощью стандартных разностных методов далеко не тривиально. Действительно, в работе [10, с. 5] утверждается, что “принцип полной консервативности . . . дает возможность строить схемы, правильно передающие быстро изменяющиеся решения даже на грубых сетках, когда фактически теряется аппроксимация”. Поскольку мы интересуемся финальной скоростью поршня, улетающего на бесконечность при  $t \rightarrow \infty$ , попробуем для расчета применить какую-либо рекомендованную в [10, с. 108] вполне консервативную разностную аппроксимацию системы уравнений газовой динамики в переменных Лагранжа с постоянным шагом  $h$  по массовой координате ( $nh = M$ ;  $u(kh, t) = u_k$ ). Тогда для любой из рекомендованных схем суммарная кинетическая энергия газа и поршня записывается в виде

$$E_{\text{kin}}(t) = 0,5h[u_1^2 + \dots + u_{n-1}^2] + 0,5(M_n + 0,5h)u_n^2. \quad (16)$$

Для потенциальной же энергии  $Q$  газа в начальный момент имеем  $Q = nh\varepsilon_0$ . Приравняв энергии  $E_{\text{kin}}(\infty) = Q$ , выводим при  $M_n = 0$  грубое неравенство  $[u_n(\infty)]^2 \leq 4n\varepsilon_0$ . Таким образом, при  $n < \gamma(\gamma-1)^{-1}$  скорость истечения газа в вакуум, равная  $[4\gamma\varepsilon_0/(\gamma-1)]^{1/2}$  (см. (9)), будет занижена в расчете в конечное число раз. Ясно, что приближение (16) для кинетической энергии, точное для нефизического “дискретного” газа с  $\gamma = \infty$ ,  $\nu = 1/2$ , становится тем грубее, чем ближе  $\gamma$  к единице, и именно потому, что в разностной схеме все больше теряется аппроксимация. Таким образом, цитированное выше утверждение из [10] несправедливо, и для эффективного решения задачи необходимо, как правило, адаптировать схему к течению. Последнее излишне лишь в случае  $M \ll \min(M_d, M_n)$ , когда вместо детального расчета можно воспользоваться классической формулой для двух масс (3).

Разумеется, для численного решения задачи Лагранжа можно успешно воспользоваться методом характеристик, адекватно отвечающим постановке зада-

чи (см. ссылки на такие расчеты в [8]). Однако, в отличие от разностных методов, метод характеристик использует далеко не шаблонную логику расчета.

Течения Лява–Пиддака–Кента принадлежат к классу одномерных автомодельных движений Седова [11], в которых скорость частицы пропорциональна расстоянию до центра симметрии газа. Среди последних есть и цилиндрические ( $\delta = 1$ ), и сферические ( $\delta = 2$ ) течения. Для них согласованная замена наружной части газа  $\{r : r \geq y\}$  на поршень массой  $M$  (или погонной массой  $M$  при  $\delta = 1$ ) задается вместо (15) равенством

$$y^{1-\delta} M = 2\delta\pi \int_y^L r\rho(r) dr, \quad dm = r^\delta \rho(r) dr. \quad (17)$$

Таким образом,  $M$  должно быть равно погонной массе убираемого газа при  $\delta = 1$ , а при  $\delta = 2$  должны совпадать их минус первые моменты [6].

Согласно [12] выход на асимптотику задает квадратура

$$UL^{-1} dt = [1 - V^{(1+\delta)(1-\gamma)}]^{-1/2} dV,$$

обобщающая (13). Поэтому при  $\delta = 2$ ,  $\gamma = 5/3$  действует тот же закон (14) разлета, что и при  $\delta = 0$ ,  $\gamma = 3$ . Поскольку (14) аналитично при  $t = 0$  и симметрично, тот же закон (14) при той же финальной скорости  $u(r) = UrL^{-1}$  и некоторой функции  $\rho(r)$  описывает при  $t \leq 0$  замедляющее изэнтропическое сжатие  $U(t) \leq 0$ , а затем при  $t \geq 0$  — последующий разлет ( $U(t) \geq 0$ ) объема газа, заключенного в соответствии с (17) внутри сферического поршня. Следует, однако, иметь в виду, что этот процесс на стадии сжатия неустойчив и вряд ли может быть реализован.

#### СПИСОК ЛИТЕРАТУРЫ

1. *Poisson S.D.* Formules relative au Mouvement du Boulet ... extraites des Manuscrite de Lagrange // J. de l'Ecole Polytechnique. Cahier 21, 1832.
2. *Love A.E.H., Pidduck P.B.* Lagrange's ballistic problem // Phil. Trans. Roy. Soc. London. Ser. A. 1922. V. 222. Pt. 1. P. 167–185. Pt. 2. P. 186–226.
3. *Станюкович К.П.* Неустановившиеся движения сплошной среды. — М.: Гостехиздат, 1955.
4. *Баум Ф.А., Станюкович К.П., Шехтер Б.М.* Физика взрыва. — М.: Физматгиз, 1959.
5. *Забабихин Е.И., Забабихин И.Е.* Явления неограниченной кумуляции. — М.: Наука, 1988.
6. *Забродин А.В., Ченцов Н.Н.* Приближенные решения в баллистической задаче Лагранжа. — М, 1988. — (Препринт / ИПМ АН СССР; № 58).
7. *Баско М.М.* О гидродинамической эффективности облучения ионными пучками. — М, 1988. — (Препринт / ИТЭФ; № 14).
8. *Корнер Дж.* Внутренняя баллистика орудий. — М.: ИЛ, 1953.
9. *Созоненко Ю.А.* Движение поршня под действием давления газа // Прикладная математика и механика. 1963. Т. 27, вып. 3. С. 535–540.
10. *Самарский А.А., Попов Ю.П.* Разностные методы решения задач газовой динамики. — М.: Наука, 1980.
11. *Седов Л.И.* Об интегрировании уравнений одномерного движения газа // Докл. АН СССР. 1953. Т. 90, № 5. С. 735.
12. *Седов Л.И.* Методы подобия и размерности в механике. — М.: Гостехиздат, 1954.