25X1

Next 1 Page(s) In Document Denied

ТЕОРИЯ ОДНОРОДНЫХ СЛУЧАЙНЫХ ПОЛЕЙ

Yaglom

А.М.Яглом (Москва, Академия наук СССР)

On the theory of homogeneous random fields

I. Введение

Центральным фактом корреляционной теории стационарных случайных процессов является факт существования спектрального разложения как для самого процесса $\xi(t)$:

(I.I) $$\xi(t)=\int_{-\infty}^{\infty} e^{it\lambda} Z(d\lambda),$$

так и для его корреляционной функции $B(\tau)=E\{\xi(t+\tau)\overline{\xi(t)}\}$:

(I.2) $$B(\tau)=\int_{-\infty}^{\infty} e^{i\tau\lambda} F(d\lambda).$$

Здесь $Z(\Lambda)$ – аддитивная случайная функция множества (случайная мера), а $F(\Lambda)$ – обычная неогрицательная ограниченная мера на $(-\infty,\infty)$, связанная с $Z(\Lambda)$ соотношением

(I.3) $$F(\Lambda)=E|Z(\Lambda)|^2.$$

Временной параметр t процесса мы предполагаем принимающим все вещественные значения; для случайных процессов с дискретным параметром пределы интегрирования в (I.I) и (I.2) надо заменить на $-\pi$ и $+\pi$.

Аналогичные спектральные разложения существуют для стационарных процессов с многомерным временным параметром $\mathbf{t}=(t_1,t_2,\dots,t_n)$ (т.е.для однородных случайных полей $\xi(\mathbf{t})$ в n-мерном пространстве R_n) и для еще более общего класса однородных полей на произвольной локально компактной коммутативной группе G (см.ниже формулы (2.2I) - (2.23)). При этом в случае однородного поля $\xi(\mathbf{t})$, $\mathbf{t}\in R_n$ любые дополнительные предположения о его симметрии могут быть сведены к специальным ограничениям на корреляционную функцию $B(\boldsymbol{\tau})$ и спектральную меру $\mathcal{F}(\Lambda)$ (или $\mathcal{Z}(\Lambda)$). Особенно интересным для приложений является случай однородного и изотропного случайного поля - однородного поля $\xi(\mathbf{t})$, обладающего сферической симметрией. Общий вид корреляционной функции $\mathcal{B}(\tau)$, $\tau = |\boldsymbol{\tau}|$, такого поля дается известной формулой Шенберга [I] :

$$(I.4) \qquad \mathcal{B}(\tau) = \int_0^\infty \frac{J_{\frac{n-2}{2}}(\tau\lambda)}{(\tau\lambda)^{\frac{n-2}{2}}} d\mathcal{G}(\lambda),$$

где $J_{\frac{n-2}{2}}$ - функция Бесселя порядка $\frac{n-2}{2}$, а $\mathcal{G}(\lambda)$ - ограниченная неубывающая функция на $[0, \infty]$). В статистическрй теории турбулентности рассматриваются также векторные однородные и изотропные случайные поля $\boldsymbol{\xi}(\mathbf{t}) = \{\xi_1(\mathbf{t}),\dots,\xi_n(\mathbf{t})\}$; отвечающая им корреляционная матрица $\mathcal{B}_{ij}(\boldsymbol{\tau}) = E\{\xi_i(\mathbf{t}+\boldsymbol{\tau})\xi_j(\mathbf{t})\}$ имеет свое "спектральное разложение", несколько напоминающее (I.4) (см. [2 - 4]). Обобщением понятия векторного изотропного случайного поля является понятие изотропного случайного потока, введенное Ито [5] ; для таких потоков также существует спектральное разложение специального вида.

Несколько в стороне от работ по теории случайных полей в евклидовых пространствах стоит небольшая заметка А.М.Обухова [6], посвященная однородным случайным полям на сфере S_2 трехмерного пространства R_3. Обухов рассмотрел разложение такого поля $\xi(\theta,\varphi)$ в ряд по сферическим функциям $Y_l^m(\theta,\varphi)$:

$$(I.5) \qquad \xi(\theta,\varphi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} z_{lm} Y_l^m(\theta,\varphi)$$

и показал,что условие однородности поля равносильно условию

$$(I.6) \qquad E z_{lm}\overline{z_{l_1m_1}} = \delta_{ll_1}\delta_{mm_1} f_l .$$

В силу теоремы сложения для присоединенных функций Лежандра отсюда для корреляционной функции $\mathcal{B}(\theta_{12}) = E\xi(\theta_1,\varphi_1)\xi(\theta_2\varphi_2)$ (где θ_{12} - угловое расстояние между точками (θ_1,φ_1) и (θ_2,φ_2) сферы S_2) получается представление

(I.7) $\mathcal{B}(\theta)=\sum_{\ell=0}^{\infty} g_\ell \mathcal{P}_\ell(\cos\theta)$, $g_\ell=\frac{2\ell+1}{2} f_\ell \geqslant 0$,

найденное в другой связи еще раньше Шенбергом[7] . Обратно, при любых $z_{\ell m}$ и f_ℓ удовлетворяющих (I.6) и таких, что ряд (I.7) сходится, поле (I.5) будет однородным, а функция (I.7) будет корреляционной функцией однородного случайного поля.

Результаты (I.2), (I.4) и (I.7) или (I.I) и (I.5) внешне представляются весьма различными. Естественно думать, однако, что все они должны содержаться в качестве частных случаев в какой-то общей теории, охватывающей широкие классы случайных функций, инвариантных относительно некоторых групп преобразований. Нетрудно понять, что единственным математическим аппаратом, пригодным для построения такой общей теории, может быть аппарат теории представлений групп. Идея настоящей работы и состоит в приложении теории представлений групп к нахождению общего вида однородных случайных полей и их корреляционных функций на различных многообразиях, допускающих транзитивную группу преобразований.

Автор искренне благодарен И.М.Гельфанду, М.А.Наймарку и Н.Я.Виленкину, беседы с которыми по вопросам теории представлений имели существенное значение для написания настоящей работы.

2. Однородные случайные поля на группах

Простейшим типом пространств, допускающих транзитивную группу преобразований, являются групповые пространства $\mathcal{G}=\{g\}$ – совокупности элементов g некоторой группы. В дальнейшем мы будем рассматривать только топологические группы $\mathcal{G}$. На группе $\mathcal{G}$ мы можем, вообще говоря, определить даже два различных семейства непрерывных преобразований: преобразования

(2.I) $V_g: \qquad g_1 \to g g_1$

(левые сдвиги) и

(2.2) $V'_g: \qquad g_1 \to g_1 g$

(правые сдвиги). Случайное поле $\xi(g)$ на группе $\mathcal{G}$ мы определим как функцию на $\mathcal{G}$ со значениями из гильбертового пространства $\mathfrak{h}$ комплекснозначных случайных величин с конечной дисперсией (скалярное произведение элементов из $\mathfrak{h}$ равно их ковариации), непрерывную в смысле сильной топологии в $\mathfrak{h}$. Поле $\xi(g_1)$ будет называться однородным, если его первые и вторые моменты не изменяются при применении к соответствующим аргументам g_1 хотя бы одного из семейств преобразований V_g и V'_g [I]. В зависимости от того, верно ли это для семейства V_g, V'_g или обоих сразу мы будем говорить о полях, однородных относительно левых сдвигов, правых сдвигов или двусторонних сдвигов. В силу транзитивности рассматриваемых групп преобразований для однородного поля всегда $E\xi(g)=\mathrm{Const}$, так что не уменьшая общности можно даже сразу считать, что $E\xi(g)\equiv 0$. При этом единственным условием однородности поля относительно левых сдвигов будет условие

(2.3) $E\xi(g_1)\overline{\xi(g_2)} = E\xi(gg_1)\overline{\xi(gg_2)} = \mathcal{B}(g_2^{-1}g_1),$

а условием однородности относительно правых сдвигов - условие

(2.4) $E\xi(g_1)\overline{\xi(g_2)} = E\xi(g_1 g)\overline{\xi(g_2 g)} = \mathcal{B}(g_1 g_2^{-1}).$

Для полей же, однородных относительно двусторонних сдвигов, должны выполняться оба условия (2.3) и (2.4) одновременно; отсюда, в частности, следует, что функция $\mathcal{B}(g_2^{-1}g_1)$ равенства (2.3) в этом случае должна сохранять свое значение при замене элементов g_1 и g_2 на $V'_g g_1$ и $V'_g g_2$, т.е. должна быть постоянной на классах сопряженных элементов:

(2.5) $\mathcal{B}(h) = \mathcal{B}(ghg^{-1}), \quad h, g \in \mathcal{G}.$

[I] Такое понятие однородности, очевидно, является понятием в широком смысле (см. [8], гл.II, § 3).

Предположим сперва, что группа $\mathcal{G}$ является компактной (в частности, она может быть просто конечной). Это предположение существенно упрощает дело, так как позволяет воспользоваться теорией представлений компактных групп, которая уже сравнительно давно достигла значительной степени завершенности (см.,например, [9] или [10]). Согласно этой теории для каждой компактной группы существует не более чем счетное число неэквивалентных между собой конечномерных унитарных неприводимых представлений (т.е.гомоморфизмов группы $\mathcal{G}$ в группу унитарных матриц конечного порядка):

(2.6) $g \to T^{(\lambda)}(g) = \| T_{ij}^{(\lambda)}(g) \|, \quad 1 \le i,j \le d_\lambda < \infty, \ \lambda = 1,2,\ldots$

(2.7) $T^{(\lambda)}(g_1 g_2) = T^{(\lambda)}(g_1) T^{(\lambda)}(g_2), \quad T^{(\lambda)}(g^{-1}) = [T^{(\lambda)}(g)]^{-1} = [T^{(\lambda)}(g)]^*.$

Матричные элементы $T_{ij}^{(\lambda)}(g)$ этих представлений удовлетворяют следующим соотношениям ортогональности

(2.8) $\int_{\mathcal{G}} T_{ij}^{(\lambda)}(g) \overline{T_{kl}^{(\mu)}(g)}\, dg = \delta_{\lambda\mu} \delta_{ik} \delta_{jl} \cdot \frac{1}{d_\lambda}$,

где $\delta_{\lambda\mu}$ - символ Кронекера, а dg - единственная инвариантная относительно всех преобразований V_g и V'_g мера на $\mathcal{G}$ такая,что $\int_{\mathcal{G}} dg = 1$. Совокупность же всех матричных элементов $T_{ij}^{(\lambda)}(g)$, $1 \le i,j \le d_\lambda$, $\lambda = 1,2,\ldots$ образует полную ортогональную систему в пространстве $L^2(\mathcal{G})$ комплексных на $\mathcal{G}$ функций с интегрируемым по dg квадратом модуля.

Пусть теперь $\xi(g)$ - произвольное случайное поле на $\mathcal{G}$. Из условий ортогональности и полноты системы функций T_{ij} следует, что $\xi(g)$ можно разложить в сходящийся в смысле квадратичного (т.е. в смысле сильной топологии в $h_{\mathcal{G}}$) ряд по этим функциям:

(2.9) $\xi(g) = \sum_\lambda \sum_{i,j=1}^{d_\lambda} z_{ji}^{(\lambda)} T_{ij}^{(\lambda)}(g)$,

где

$$(2.10)\quad z_{ji}^{(\lambda)} = d_\lambda \int_{\mathcal{G}} \xi(g)\overline{T_{ij}^{(\lambda)}(g)}\,dg \in h_{\mathcal{Y}}$$

(интеграл в (2.10) также понимается в смысле среднего квадратичного). Для полей с $E\xi(g)\equiv 0$, очевидно, $Ez_{ji}^{(\lambda)}\equiv 0$. Если предположить далее, что поле $\xi(g)$ однородно, то это, естественно, налагает дальнейшие существенные ограничения на коэффициенты $z_{ji}^{(\lambda)}$.

Теорема I. Для того, чтобы случайное поле (2.9) было однородным относительно левых сдвигов, необходимо и достаточно, чтобы случайные величины $z_{ji}^{(\lambda)}$ удовлетворяли условиям

$$(2.11)\quad E z_{ji}^{(\lambda)}\overline{z_{\ell k}^{(\mu)}} = \delta_{\lambda\mu}\delta_{ik} f_{j\ell}^{(\lambda)},$$

где матрицы

$$(2.12)\quad f^{(\lambda)} = \| f_{j\ell}^{(\lambda)} \|$$

являются положительно определенными и такими, что

$$(2.13)\quad \sum_\lambda \sum_j f_{jj}^{(\lambda)} = \sum_\lambda Tr\, f^{(\lambda)} < \infty$$

Корреляционная функция $\mathcal{B}(g)$ равенства (2.3) в этом случае будет представляться в виде

$$(2.14)\quad \mathcal{B}(g) = \sum_\lambda \sum_{j,\ell} f_{j\ell}^{(\lambda)} T_{\ell j}^{(\lambda)}(g).$$

Обратно, любая функция $\mathcal{B}(g)$ вида (2.14), где $\| f_{j\ell}^{(\lambda)} \|$ - положительно определенные матрицы, удовлетворяющие (2.13) будет корреляционной функцией некоторого однородного относительно левых сдвигов случайного поля из $\mathcal{G}$.

Доказательство. Из равенств (2.3), (2.10), (2.7), (2.8) и того, что $dg_2^{-1}g = dg$ легко следует соотношение

$$(2.15)\quad E z_{ji}^{(\lambda)}\overline{z_{\ell k}^{(\mu)}} = \delta_{\lambda\mu}\delta_{ik} d_\lambda \int_{\mathcal{G}} \mathcal{B}(g) T_{\ell j}^{(\lambda)}(g)\,dg = \delta_{\lambda\mu}\delta_{j\ell}\, f_{j\ell}^{(\lambda)}$$

Тем самым доказана необходимость условия (2.11). Подставив затем (2.9) в (2.3), убеждаемся в справедливости (2.14). Матрицы $f^{(\lambda)}$

очевидно все будут неотрицательно определенными; так как $T^{(\lambda)}_{ik}(e) = \delta_{ik}$ (где e – единица группы $\mathcal{G}$), то условие (2.I3) будет условием сходимости рядов в правых частях (2.9) и (2.I4). Если теперь $f^{(\lambda)}$, $\lambda = 1, 2, \ldots$ – произвольные неотрицательно определенные $(d_\lambda \times d_\lambda)$-матрицы, то всегда можно подобрать $z_{ji} \in \mathfrak{h}$ так, чтобы выполнялось (2.II). При условии (2.I3) соответствующий ряд в правой части (2.9) будет сходиться и будет определять случайное поле $\xi(g)$, для которого $E\xi(g_1 g)\overline{\xi(g_1)}$ будет даваться формулой (2.I4). Следовательно, поле $\xi(g)$ будет однородным относительно левых сдвигов, что и завершает доказательство теоремы I.

Хорошо известно, что класс корреляционных функций случайного поля совпадает с классом положительно определенных функций в соответствующем пространстве (ср., например, [8] , глава П, теорема 3.I). Таким образом, формула (2.I4) определяет общий вид функции $\mathcal{B}(g)$ на $\mathcal{G}$, таких, что при любых комплексных $a_1, a_2, \ldots, a_n$

$$\sum_{i,k} \mathcal{B}(g_k^{-1} g_i)\, a_i \overline{a_k} \geqslant 0. \tag{2.I6}$$

Этот последний результат был еще сравнительно давно получен Бохнером [II] . Отметим, что все остальные утверждения теоремы I можно получить отсюда, если воспользоваться общей теоремой Карунена [I2] и Крамера [I3] о разложении случайных функций. Однако, приведенное выше доказательство теоремы I является более прямым и элементарным.

Совершенно аналогично обстоит дело и в случае полей на $\mathcal{G}$, однородных относительно правых сдвигов; при этом только условие (2.II) заменяется условием

$$E z^{(\lambda)}_{ji} \overline{z^{(\mu)}_{\ell k}} = \delta_{\lambda\mu}\, \delta_{j\ell}\, f^{(\lambda)}_{ik} \tag{2.I7}$$

В случае же полей $\xi(g)$, однородных относительно двусторонних сдвигов, должны выполняться оба равенства (2.II) и (2.I7), так что имеет место

Теорема 2. Для того, чтобы случайное поле (2.9) было однородным относительно двусторонних сдвигов, необходимо и достаточно, чтобы случайные величины $z_{ji}^{(\lambda)}$ удовлетворяли соотношению:

$$E z_{ji}^{(\lambda)} \overline{z_{\ell k}^{(\mu)}} = \delta_{\lambda\mu} \delta_{ik} \delta_{j\ell} f^{(\lambda)}, \tag{2.18}$$

где $f^{(\lambda)}$ - неотрицательные числа такие, что

$$\sum_{\lambda} d_{\lambda} f^{(\lambda)} < \infty \tag{2.19}$$

Корреляционная функция $\mathcal{B}(g)$ равенства (2.3) при этом будет представляться в виде

$$\mathcal{B}(g) = \sum_{\lambda} f^{(\lambda)} \chi^{(\lambda)}(g), \tag{2.20}$$

где $\chi^{(\lambda)}(g) = Tr\left[T^{(\lambda)}(g)\right]$ - характеры группы $\mathcal{G}$.

Формула (2.20) для общей положительно определенной функции на компактной группе $\mathcal{G}$, инвариантной относительно двусторонних сдвигов, была впервые указана С.Бохнером [II].

Теорема 2 применима, в частности, в случае абелевой группы $\mathcal{G}$ (на которой, разумеется, существуют лишь двусторонние однородные поля). В этом частном случае предположение о компактности группы $\mathcal{G}$ оказывается излишним; его вполне можно заменить, например, предположением о локальной компактности (или даже некоторым еще более общим предположением; ср. ([I4]). Общий вид положительно определенной функции на $\mathcal{G}$ при этом дается следующим известным обобщением классической теоремы Бохнера:

$$\mathcal{B}(g) = \int_{\overline{\mathcal{G}}} \chi^{(\lambda)}(g)\, \mathcal{F}(d\lambda), \tag{2.21}$$

где $\mathcal{F}(d\lambda)$ - ограниченная мера на множестве характеров $\overline{\mathcal{G}}$ группы $\mathcal{G}$

(Вейль [I0] , Райков [I4]). В силу теоремы о разложении Карунена-Крамера отсюда немедленно следует, что любое однородное поле $\xi(g)$ на такой группе допускает спектральное разложение вида

$$(2.22) \qquad \xi(g) = \int_{\mathcal{G}} \chi^{(\lambda)}(g) Z(d\lambda),$$

где $Z(d\lambda)$ – случайная мера на $\mathcal{G}$ такая,что

$$(2.23) \qquad E Z(\Lambda_1)\overline{Z(\Lambda_2)} = \mathcal{F}(\Lambda_1 \cap \Lambda_2)$$

(см. [I5] , [I6]). Обычное спектральное разложение (I.I)-(I.2) стационарных процессов $\xi(t)$, очевидно, является частным случаем разложения (2.2I)-(2.23).

Значительно сложнее обстоит дело в случае произвольной локально компактной группы $\mathcal{G}$ (не предполагаемой коммутативной). Такая группа, как известно, может вообще не иметь ни одного конечномерного унитарного представления. Правда, как было показано Гельфандом и Райковым [I7] она всегда будет иметь достаточное количество бесконечномерных унитарных представлений – гомеоморфных отображений $\mathcal{G}$ в множество унитарных операторов гильбертового пространства. Однако вопрос о разложении произвольной функции $\xi(g)$ на группе $\mathcal{G}$ по матричным элементам таких представлений даже и для обычных (не случайных) функций пока еще не разобран в достаточно общей форме; поэтому вряд ли можно рассчитывать на непосредственное перенесение метода доказательства теоремы I на сравнительно широкие классы локально-компактных групп. Значительно более перспективным представляется путь, исходящий из построения аналога теоремы Бохнера, описывающего всевозможные положительно определенные функции на $\mathcal{G}$.

Как показано в работе [I7] (см.также книгу [I8]) каждая положительно определенная функция $B(g)$ на локально компактной группе $\mathcal{G}$ представима в виде

(2.24) $\mathcal{B}(g)=(T(g)\xi_0,\xi_0)$,

где ξ_0 - определенный вектор гильбертова пространства H, в котором действует некоторое унитарное представление группы $\mathcal{G}$, а $T(g)$ - оператор представления. Таким образом, описание всевозможных положительно определенных функций сводится к описанию всевозможных унитарных представлений $\mathcal{G}$. Каждое такое представление может быть разложено в прямой интеграл неприводимых представлений (см., например, [18], глава УШ), однако, в общем случае такое разложение будет в высокой степени неоднозначным и не может быть использовано для получения каких-либо более или менее окончательных формул[1]. Значительно более удовлетворительные результаты удается получить, если наложить некоторые ограничены на класс рассматриваемых групп $\mathcal{G}$. А именно, мы будем считать, что группа $\mathcal{G}$ является сепарабельной группой типа I, т.е. такой сепарабельной группой, что каждое ее унитарное представление порождает кольцо операторов, являющееся кольцом типа I в смысле Меррея и фон Неймана (ср. [18], глава УП). Согласно результатам Хариш-Чандра [19] этому условию будут удовлетворять, в частности, все связные полупростые группы Ли; повидимому, оно будет выполняться и для большинства других "достаточно хороших" локально компактных групп (ср., например, [20]). В то же время указанное условие поз-

---

[1] Исходя из такого разложения для положительно определенной функции $\mathcal{B}(g)$ можно получить лишь формулу вида

$$\mathcal{B}(g)=\int_S Tr\left(T^{(\lambda)}(g)A(d\lambda)\right),$$

где S - какое-то пространство с мерой $d\lambda$, $T^{(\lambda)}(g)$ - зависящее от $\lambda\in S$ непроводимое унитарное представление группы $\mathcal{G}$ и $A(d\lambda)$ "операторная мера" на S (ср. ниже стр. 11).

воляет воспользоваться результатами работ Макки [2I] и Гишарде [22], согласно которым любое унитарное представление сепарабельной группы $\mathcal{G}$ типа I всегда может быть представлено в виде прямой суммы однократных представлений, разлагающихся в интеграл по неприводимым неэквивалентным представлениям. Точнее говоря, пространства H_ℓ, в которых действуют эти однократные представления $T_\ell(g)$ можно представить как топологические прямые интегралы гильбертовых пространств $H^{(\lambda)}$, в которых действуют неэквивалентные между собой неприводимые унитарные представления $T^{(\lambda)}(g)$ группы $\mathcal{G}$:

(2.25) $$T_\ell(g) = \int_{\bar{\mathcal{G}}} \oplus T^{(\lambda)}(g)\,, \quad H_\ell = \int_{\bar{\mathcal{G}}} H^{(\lambda)} \sqrt{\sigma_\ell(d\lambda)}$$

(ср. [I8], глава УШ). Здесь $\bar{\mathcal{G}}$ означает "дуальный объект" группы $\mathcal{G}$ - совокупность всевозможных классов неприводимых неэквивалентных унитарных представлений этой группы, а через $\sigma_\ell(d\lambda)$ обозначена некоторая мера в пространстве $\bar{\mathcal{G}}$ (снабженном "естественной борелевской структурой" [2I] - борелевским полем измеримых множеств); представление $T_\ell(g)$ при этом естественно зависит лишь от класса эквивалентных мер в $\bar{\mathcal{G}}$, к которому принадлежит σ_ℓ. Если теперь мы подставим в (2.24) вместо $T(g)$ прямую сумму представлений (2.25), то для $\mathcal{B}(g)$ немедленно получится формула вида

$$\mathcal{B}(g) = \int_{\bar{\mathcal{G}}} \sum_{i,j,\ell} T_{ij}^{(\lambda)}(g)\, \xi_{i,\ell}^{(\lambda)} \overline{\xi_{j,\ell}^{(\lambda)}}\, \sigma^{(\ell)}(d\lambda)\,,$$

которую можно также переписать в виде

(2.26) $$\mathcal{B}(g) = \int_{\bar{\mathcal{G}}} \mathrm{Tr}\left(T^{(\lambda)}(g)\,\mathcal{F}(d\lambda)\right),$$

где $\mathcal{F}(d\lambda)$ - некоторый аддитивно зависящий от $d\lambda$ эрмитовый неотрицательно определенный оператор в пространстве $H^{(\lambda)}$ представлений $T^{(\lambda)}(g)$ [I].

Таким образом мы приходим к следующей теореме:

Теорема 3. Если $\mathcal{G}$ - сепарабельная локально компактная группа типа I, то класс положительно определенных функций на $\mathcal{G}$ совпадает с классом функций, представимых в виде (2.26), где $\mathcal{F}(d\lambda)$ - "операторная мера" на $\bar{\mathcal{G}}$ (значения которой являются эрмитовыми неотрицательными операторами в гильбертовом пространстве $H^{(\lambda)}$) такая, что

$$\int_{\bar{\mathcal{G}}} Tr\, \mathcal{F}(d\lambda) = Tr\, \mathcal{F}(\bar{\mathcal{G}}) < \infty . \tag{2.27}$$

В самом деле, то что любая функция вида (2.26) является положительно определенной, легко проверяется непосредственно; условие же (2.27), очевидно, гарантирует сходимость интеграла (2.26) Формула (2.26) является естественным обобщением соотношения (2.I4) (в компактном случае все условия теоремы 3 выполняются очевидным образом).

Пусть теперь $\xi(g)$ - однородное относительно левых сдвигов случайное поле на группе $\mathcal{G}$, удовлетворяющей всем условиям теоремы 3, а функция (2.26) - корреляционная функция этого поля. Для получения "спектрального разложения" поля $\xi(g)$ можно воспользоваться методом, аналогичным примененному в работах [I2] и [I3]. Рассмотрим гильбертово пространство $L^2(\mathcal{F})$ всевозможных операторных функций $U^{(\lambda)}$, $\lambda \in \bar{\mathcal{G}}$ со значениями из кольца ограниченных

I Если среди унитарных представлений группы $\mathcal{G}$ имеются и конечномерные, то интегралы (2.25) и (2.26) и все последующие интегралы по $\bar{\mathcal{G}}$ естественно распадаются на сумму интегралов по подпространствам $\bar{\mathcal{G}}_n \subset \bar{\mathcal{G}}$, $n = 1, 2, \dots, \infty$, отвечающим всевозможным неэквивалентным n-мерным неприводимым унитарным представлениям группы $\mathcal{G}$. В каждом же слагаемом такой суммы пространства $H^{(\lambda)}$ (имеющие одинаковую размерность) все считаются совпадающими.

операторов в $H^{(\lambda)}$ и нормой

(2.28) $\| \mathcal{U}^{(\lambda)} \|^2 = \int_{\bar{\mathcal{G}}} \mathrm{Tr}\left[\mathcal{U}^{(\lambda)} \mathcal{F}(d\lambda) \mathcal{U}^{(\lambda)*} \right] < \infty$.

В таком случае отображение

(2.29) $T^{(\lambda)}(g) \leftrightarrow \xi(g)$

будет изометрическим отображением множества случайных величин $\{\xi(g),\, g \in \mathcal{G}\}$ в $L^2(\mathcal{F})$, которое можно продолжить до линейного изометрического отображения всего пространства $L^2(\mathcal{F})$ в $\mathfrak{H}$ (ср. [I2]) Пусть теперь f_1 и f_2 – два произвольных вектора пространства $H^{(\lambda)}$, а Λ – измеримое множество в $\bar{\mathcal{G}}$; тогда тройке (Λ, f_1, f_2) мы следующим образом сопоставим операторную функцию $\mathcal{U}^{(\lambda)}(\Lambda; f_1, f_2) \in L^2(\mathcal{F})$:

(2.30) $\mathcal{U}^{(\lambda)}(\Lambda; f_1, f_2) f = \begin{cases} (f, f_2) f_1 , & \text{если } \lambda \in \Lambda \\ 0 , & \text{если } \lambda \bar{\in} \Lambda \end{cases}$

Пусть $\mathcal{Z}(\Lambda; f_1, f_2)$ – образ $\mathcal{U}^{(\lambda)}(\Lambda; f_1, f_2)$ при изометрическом отображении (2.29). В таком случае $\mathcal{Z}(\Lambda; f_1, f_2)$ – случайная величина билинейно зависящая от f_1 и f_2 ; ее можно записать в виде

(2.3I) $\mathcal{Z}(\Lambda; f_1, f_2) = (\mathcal{Z}(\Lambda) f_1, f_2)$,

где $\mathcal{Z}(\Lambda)$ – случайный оператор в $H^{(\lambda)}$, аддитивно зависящий от $\Lambda \subset \bar{\mathcal{G}}$. Пусть $e_1, e_2, \ldots, e_n \ldots$ – произвольный ортонормированный базис в $H^{(\lambda)}$, тогда нетрудно проверить, что в $L^2(\mathcal{F})$ операторная функция

$$\int_{\bar{\mathcal{G}}} \sum_{i,j} (T^{(\mu)}(g) e_i, e_j)\, \mathcal{U}^{(\lambda)}(d\mu; e_j, e_i)$$

будет совпадать с $T^{(\lambda)}(g)$. Отсюда для $\xi(g)$ – образа $T^{(\lambda)}(g)$ при нашем изометрическом отображении – немедленно получается формула

(2.32) $\xi(g) = \int_{\bar{\mathcal{G}}} \mathrm{Tr}\left[T^{(\lambda)}(g) \mathcal{Z}(d\lambda) \right]$.

Кроме того в силу изометричности этого отображения

$$E(\mathcal{Z}(\Lambda_1) f_1, f_2)\overline{(\mathcal{Z}(\Lambda_2) g_1, g_2)} = \int_{\bar{\mathcal{G}}} \mathrm{Tr}\left[\mathcal{U}^{(\lambda)}(\Lambda_2; f_1, f_2) \mathcal{F}(d\lambda) \mathcal{U}^{(\lambda)*}(\Lambda_1; g_1, g_2) \right].$$

т.е.

$$(2.33)\quad E\big(\mathcal{Z}(\Lambda_1)f_1,f_2\big)\overline{\big(\mathcal{Z}(\Lambda_2)g_1,g_2\big)}=(f_1,g_1)\big(\mathcal{F}(\Lambda_1\cap\Lambda_2)g_2,f_2\big).$$

Обратно, для любого случайного поля $\xi(g)$ вида (2.32), где $\mathcal{Z}(\Lambda)$ - случаыйный линейный оператор в $H^{(\lambda)}$, удовлетворяющий (2.33), $E\xi(gg_1)\overline{\xi(g_1)}$ будет, как легко видеть, равно правой части (2.26), т.е. $\xi(g)$ будет однородно относительно правых сдвигов. Таким образом, имеет место

Теорема 4. Для того, чтобы случайное поле $\xi(g)$ на сепарабельной локально компактной группе $\mathcal{G}$ типа I было однородным относительно левых сдвигов, необходимо и достаточно, чтобы оно могло быть представлено в виде (2.32), где $\mathcal{Z}(\Lambda)$ - случайный оператор в $H^{(\lambda)}$, аддитивно зависящий от множества $\Lambda\subset\bar{\mathcal{G}}$ и удовлетворяющий при любых $f_1,f_2,g_1,g_2\in H^{(\lambda)}$ условию (2.33). В этом условии $\mathcal{F}(\Lambda)$ - это эрмитов неотрицательный оператор в $H^{(\lambda)}$, удовлетворяющий (2.27), через который корреляционная функция $\mathcal{B}(g)$ поля $\xi(g)$ выражается согласно формуле (2.26).

Теорему 4 можно рассматривать как обобщение теоремы I на случай локально компактных групп. В случае полей однородных относительно правых вдсигов в условиях теоремы 4 надо лишь заменить условие (2.33) на условие

$$(2.34)\quad E\big(\mathcal{Z}(\Lambda_1)f_1,f_2\big)\overline{\big(\mathcal{Z}(\Lambda_2)g_1,g_2\big)}=\big(\mathcal{F}(\Lambda_1\cap\Lambda_2)f_1,g_1\big)(g_2,f_2).$$

Если мы выберем в $H^{(\lambda)}$ определенную "систему координат" (ортонормированный базис), то формулы (2.26), (2.27) и (2.28), (2.33) можно будет переписать в виде весьма близком к (2.I0)-(2.I4)

$$(2.26')\quad \mathcal{B}(g)=\int_{\bar{\mathcal{G}}}\sum_{i,j}T_{ij}^{(\lambda)}(g)\mathcal{F}_{ji}(d\lambda),\quad \sum_i\mathcal{F}_{ii}(\bar{\mathcal{G}})<\infty,$$

$$(2.32')\quad \xi(g)=\int_{\bar{\mathcal{G}}}\sum_{i,j}T_{ij}^{(\lambda)}(g)\mathcal{Z}_{ji}(d\lambda),$$

где в случае полей, однородных относительно левых сдвигов,

(2.33') $E Z_{ij}(\Lambda_1)\overline{Z_{kl}(\Lambda_2)} = \delta_{jl} F_{ik}(\Lambda_1 \cap \Lambda_2)$,

а в случае однородности относительно правых сдвигов

(2.34') $E Z_{ij}(\Lambda_1)\overline{Z_{kl}(\Lambda_2)} = \delta_{ik} F_{jl}(\Lambda_1 \cap \Lambda_2)$.

Для полей, однородных относительно двусторонних сдвигов, должны выполняться и соотношения (2.33) и соотношения (2.34). Отсюда вытекает, что корреляционная функция $B(g)$ здеть должна разлагаться по следам операторов представления $T^{(\lambda)}(g)$. Так как бесконечномерные унитарные операторы не имеют конечных следов, то отсюда следует, что некомпактных некоммутативных группах, вообще говоря, может и не существовать случайных полей, однородных относительно двусторонних сдвигов (см., впрочем, ниже стр.).

В заключение приведем некоторые ссылки на работы, содержащие явные формулы для операторов представлений $T^{(\lambda)}(g)$ на тех или иных группах; исходя из этих формул, очевидно, можно явно перечислить также и все однородные поля на соответствующих группах. Для группы вращений трехмерного эвклидова пространства значения всех матричных элементов $T_{ij}^{(\lambda)}(g)$ (в функциях от трех углов Эйлера, однозначно задающих вращение g) выписаны в книге [23] ; в общем случае эти матричные элементы выражаются через тригонометрические функции и полиномы Якоби. Для группы движений евклидовой плоскости (имеющей однопараметрическое семейство бесконечномерных унитарных представлений) явный вид матричных элементов $T_{ij}^{(\lambda)}(g)$ указан в работе [24] (эти матричные элементы выражаются через тригонометрические и бесселевы функции). Для группы движений плоскости Лобачевского матричные элементы $T_{ij}^{(\lambda)}(g)$ выписаны в работе [25] - здесь они выражаются через гипергеометрические функции. В общем случае произвольной классической группы явные выражения для операторов $T^{(\lambda)}(g)$ (задаваемых не в матричной форме) можно найти в книге [26] . Методы построения формул для всех матричных элементов представлений группы вращений и группы движений евклидова n-мерного пространства и группы движений n-мерного пространства Лобачевского указаны в заметках [27], [28] и [29] (заметим, что эти матричные элементы выражаются через некоторые новые, до сих пор не изученные, трансцендентные функции).

3. Однородные поля на однородных пространствах.

Пусть $X=\{x\}$ – произвольное однородное пространство, т.е. пространство, допускающее транзитивную группу преобразований $\mathcal{G}=\{g\}$. Обозначим через $\mathcal{K}=\{k\}$ стационарную подгруппу $\mathcal{G}$, т.е. подгруппу оставляющую на месте некоторую точку $x_0 \in X$. Совокупность преобразований $g \in \mathcal{G}$, переводящих x_0 в фиксированную точку $x_1 \in X$, образует, очевидно, левый класс смежности $g_1\mathcal{K}$ группы $\mathcal{G}$ по $\mathcal{K}$. Тем самым между точками $x_1 \in X$ и левыми классами смежности $g_1\mathcal{K}$ устанавливается взаимно однозначное соответствие, так что X можно отождествить с множеством этих классов смежности: $X=\mathcal{G}/\mathcal{K}$; если $x_1=g_1\mathcal{K}$, то $gx_1=gg_1\mathcal{K}$. В дальнейшем мы всегда будем предполагать, что вруппа $\mathcal{G}$ топологическая, а $\mathcal{K}$ – ее замкнутая компактная подгруппа. В частном случае, когда $\mathcal{K}$ – единичная подгруппа, однородное пространство X обращается в групповое пространство $\mathcal{G}=\{g\}$.

Топология группового пространства $\mathcal{G}$ естественно индуцирует топологию в X (см. [10]). При этом функции на X оказываются непрерывными тогда и только тогда, когда непрерывны соответствующие функции на $\mathcal{G}$ (принимающие постоянные значения на всех левых классах смежности $\mathcal{G}$ по $\mathcal{K}$). Случайное поле $\xi(x)$ на X мы определим, как непрерывное отображение X в $\mathfrak{H}$. Поле $\xi(x)$ называется однородным, если его первые и вторые моменты не меняются при преобразованиях $g \in \mathcal{G}$, т.е. если $E\xi(x)$ равно постоянной (которую мы всегда будем считать равной нулю), а $\mathcal{B}(x_1,x_2) = E\xi(x_1)\overline{\xi(x_2)}$ удовлетворяет соотношению:

$$(3.1) \quad \mathcal{B}(x_1,x_2) = \mathcal{B}(gx_1,gx_2), \quad g \in \mathcal{G}.$$

Очевидно, что класс однородных случайных полей на X, совпадает с классом однородных случайных полей на $\mathcal{G}$, постоянных на всех левых классах смежности по $\mathcal{K}$.

Начнем опять с простейшего случая, когда группа $\mathcal{G}$ компактна. Здесь можно воспользоваться общей теорией сферических функций (сферических гармоник) на компактных однородных пространствах, развитой Э.Картаном [30] и Г.Вейлем [3I] (см.также [I0]). Рассмотрим полную систему унитарных неприводимых неэквивалентных представлений (2.6) группы $\mathcal{G}$ и выберем в пространствах этих представлений базис так,чтобы они распались на неприводимые представления подгруппы $\mathcal{K}$. Для того,чтобы матричный элемент $T_{ij}^{(\lambda)}(g)$ был постоянным на всех левых классах смежности $\mathcal{G}$ по $\mathcal{K}$ должно выполняться равенство

$$(3.2) \quad T_{ij}^{(\lambda)}(gk) \equiv \sum_m T_{im}^{(\lambda)}(g)T_{mj}^{(\lambda)}(k) = T_{ij}^{(\lambda)}(g), \quad g \in \mathcal{G}; \ k \in \mathcal{K},$$

$$(3.3) \quad T_{mj}^{(\lambda)}(k) = \delta_{mj}, \quad m = 1,2,\dots,d_\lambda, \quad k \in \mathcal{K}.$$

Отсюда видно, что постоянные на левых классах смежности матричные элементы $T_{ij}^{(\lambda)}(g)$ заполняют столбцы матрицы $T^{(\lambda)}(g)$, отвечающие единичным представлениям $\mathcal{K}$. Пусть,например, представление $T^{(\lambda)}$ группы $\mathcal{G}$ равно r_λ раз содержит единичное представление $\mathcal{K}$. Предположим,что в нашем базисе $e_1, e_2, \dots, e_{d_\lambda}$ эти единичные представления отвечают первым r_λ базисным векторам (так что $T^{(\lambda)}(k) \cdot e_j = e_j$ при $k \in \mathcal{K}$ и $j = 1,\dots,r_\lambda$). В таком случае функции от x

(3.4) $\Phi_{ij}^{(\lambda)}(x) = T_{ij}^{(\lambda)}(g), \quad i=1,\dots,d_\lambda, \; j=1,\dots,r_\lambda, \; \lambda=1,2,\dots$

будут называться сферическими функциями на X, а функции

(3.5) $\Phi_{ij}^{(\lambda)}(x) = T_{ij}^{(\lambda)}(g), \quad i=1,\dots,r_\lambda, \; j=1,\dots,r_\lambda, \; \lambda=1,2,\dots$

- зональными сферическими функциями. Легко видеть, что зональные сферические функции принимают постоянные значения на всех двусторонних классах смежности $\mathcal{K}g\mathcal{K}$ группы $\mathcal{G}$ по $\mathcal{K}$; иначе говоря, для них $\Phi_{ij}^{(\lambda)}(x) = \Phi_{ij}^{(\lambda)}(kx)$, $k\in\mathcal{K}$. Совокупность точек kx, $k\in\mathcal{K}$, естественно назвать сферой с центром в точке $x_0 (=\mathcal{K})$, проходящей через точку x; таким образом, функции (3.5) постоянны на всех сферах с центром в x_0. Поэтому зональная функция $\Phi_{ij}^{(\lambda)}(x)$ фактически будет зависеть лишь от сововкупности инвариантов упорядоченной пары точек x и x_0, сохраняющихся при применении к обоим точкам любого преобразования $g\in\mathcal{G}$ (от сложного расстояния от x до x_0):

(3.6) $\Phi_{ij}^{(\lambda)}(x) \equiv \Phi_{ij}^{(\lambda)}(x, x_0) = \Phi_{ij}^{(\lambda)}(gx, gx_0), \quad g\in\mathcal{G}, \; i,j=1,\dots,r_\lambda, \; \lambda=1,2,\dots$

Согласно общей теории сферических функций функции (3.4) представляют собой полную ортогональную систему в пространстве $L^2(X)$ функций на X с квадратом модуля, интегрируемым по мере dx, инвариантной относительно всех преобразований $g\in\mathcal{G}$. Только эти функции (3.4) и входят в разложение функции $\varphi(g)$, постоянной на всех левых классах смежности $\mathcal{G}/\mathcal{K}$, по матричным элементам $T_{ij}^{(\lambda)}(g)$ ([10], гл.У, § 23).

Применение этой теории к однородным случайным полям $\xi(x)$ в силу теоремы I приводит к следующему результату:

Теорема 5. Для того, чтобы случайное поле

(3.7) $\xi(x) = \sum_\lambda \sum_{i=1}^{d_\lambda} \sum_{j=1}^{r_\lambda} z_{ji}^{(\lambda)} \Phi_{ij}^{(\lambda)}(x), \quad z_{ji}^{(\lambda)} = \int_X \xi(x)\overline{\Phi_{ij}^{(\lambda)}(x)}\,dx \Big/ \int_X |\Phi_{ij}^{(\lambda)}(x)|^2 dx$

на компактном однородном пространстве $X = \mathcal{G}/\mathcal{K}$ было однородным необходимо и достаточно, чтобы случайные величины $z_{ji}^{(\lambda)}$ удовлетворяли соотношениям

$$(3.8)\qquad E z_{ji}^{(\lambda)} \overline{z_{\ell k}^{(\mu)}} = \delta_{\lambda\mu}\delta_{ik} f_{j\ell}^{(\lambda)}.$$

При этом корреляционная функция $\mathcal{B}(x_1, x_2)$ поля $\xi(g)$ будет представляться в виде

$$(3.9)\qquad \mathcal{B}(x_1, x_2) = \sum_{\lambda} \sum_{j,\ell=1}^{r_\lambda} f_{j\ell}^{(\lambda)} \Phi_{\ell j}^{(\lambda)}(x_1, x_2),$$

где $\Phi_{\ell j}(x_1, x_2)$ – функции (3.6). Обратно, любая функция $\mathcal{B}(x_1, x_2)$ вида (3.9), где $\| f_{j\ell}^{(\lambda)} \|$ – эрмитово неотрицательно определенные матрицы такие, что ряд (3.9) сходится, будет корреляционной функцией некоторого однородного поля на X.

Формула (3.9) представляет собой несколько усовершенствованную запись результата, сформулированного еще в I94I г. Бохнером [II]. Частным случаем теоремы 5 является теорема Обухова [6] об однородных полях на двумерной сфере S_2, упоминавшаяся во введении к настоящей статье. В более общем случае однородных полей на сфере S_{n-1} n-мерного евклидова пространства наша теорема означает, что каждое такое поле будет разлагаться в ряд по гиперсферическим гармоникам $Y_{\ell, m_1, \dots, m_{n-3}, \pm m_{n-2}}(\theta_1, \dots, \theta_{n-2}, \varphi)$, $\ell = 0, 1, 2, \dots, 0 \le m_{n-2} \le m_{n-3} \le \dots \le m_1 \le \ell$ (выражающимся через тригонометрические функции и функции Гегенбауэра от сферических координат $\theta_1, \dots, \theta_{n-2}, \varphi$; см.[32], т.П) с некоррелированными коэффициентами $z_{\ell, m_1, \dots, m_{n-3}, \pm m_{n-2}}$, дисперсия которых зависит лишь от ℓ. Зональные сферические функции в этом случае будут даваться полиномами Гегенбауэра ("ультрасферическими полиномами") $C^{(\frac{n-2}{2})}(\mu)$ от $\mu = \cos\theta$; поэтому формула (3.9) здесь обращается в известную формулу Шенберга [7] для положительно определенных функций на $(m-1)$-мерной сфере:

$$(3.10)\qquad \mathcal{B}(\theta_{12}) = \sum_{\ell=0}^{\infty} f_\ell\, C_\ell^{(\frac{n-2}{2})}(\cos\theta_{12})$$

(θ_{12} - угловое расстояние между точками x_1 и x_2 сферы S_{m-1}). То обстоятельство, что в этом случае сложное расстояние θ_{12} (которое здесь задается одним числом) симметрично зависит от точек x_1 и x_2 (так что можно говорить просто о расстоянии между двумя точками) и то что $\tau_\lambda \leq 1$ при любом λ имеет общее объяснение, которое будет ясно из дальнейшего.

Перейдем теперь к случаю локально компактных однородных пространств X. Начнем с задачи об определении общего вида положительно определенной функции $\mathcal{B}(x_1, x_2)$ на X, удовлетворяющей условию (3.I). Каждой такой функции можно следующим образом однозначно сопоставить положительно определенную функцию $\mathcal{B}(g)$ на $\mathcal{G}$:

(3.II) $\mathcal{B}(g) = \mathcal{B}(x_1, x_2)$, если $g = k_1 g_2^{-1} g_1 k_2$, $x_1 = g_1 \mathcal{K}$, $x_2 = g_2 \mathcal{K}$.

Очевидно, что при этом $\mathcal{B}(g) = \mathcal{B}(k_1 g k_2)$, при любых $k_1, k_2 \in \mathcal{K}$, т.е. $\mathcal{B}(g)$ принимает постоянные значения на всех двусторонних классах смежности $\mathcal{G}$ по $\mathcal{K}$. Обратно, положительно определенной функции на $\mathcal{G}$, постоянной на всех двусторонних классах смежности по $\mathcal{K}$, мы можем при помощи равенства (3.II) сопоставить положительно определенную функцию на X, удовлетворяющую (3.I). Таким образом наша задача сводится к разысканию всех положительно определенных функций на $\mathcal{G}$, постоянных на двусторонних классах смежности по $\mathcal{K}$.

В силу [I7] любая положительно определенная функция на задается формулой (2.24). Для того, чтобы она была постоянной на всех двусторонних классах смежности по $\mathcal{K}$ необходимо и достаточно чтобы вектор ξ_0 удовлетворял условию:

(3.I2) $T(k)\xi_0 = \xi_0$ при всех $k \in \mathcal{K}$.

В частном случае, когда унитарное представление $T(g)$ группы $\mathcal{G}$ является неприводимым, функция (2.24), где ξ_0 удовлетворяет (3.I2) будет зональной сферической функцией, отвечающей этому представлению. Вообще, функция $\Phi(x)$ называется зональной сферической функцией на X, отвечающей неприводимому унитарному представлению $T^{(\lambda)}(g)$, если она представима в виде

(3.I3) $\Phi(x)=\Phi(x_0,x)=(T^{(\lambda)}(g)\xi,\eta),\ T^{(\lambda)}(k)\xi=\xi, T^{(\lambda)}(k)\eta=\eta$ при всех $k\in\mathcal{K}$

и называется просто сферической функцией, если

(3.I4) $\Phi(x)=(T^{(\lambda)}(g)\xi,\eta),\quad T^{(\lambda)}(k)\xi=\xi$ при всех $k\in\mathcal{K}$.

Очевидно, что зональные сферические функции (3.I3) зависят лишь от сложного расстояния от x до x_0, а функции (3.I4) – от точки $x=g\mathcal{K}\in X$.

Предположим теперь, что группа $\mathcal{G}$ – сепарабельная локально компактная группа типа I. В таком случае в силу [2I], [22] мы можем разложить представление $T(g)$ формулы (2.24) в прямую сумму однократных представлений, каждое из которых в свою очередь разлагается в континуальную прямую сумму неприводимых неэквивалентных представлений. Если ξ_0 удовлетворяет (3.I2), то проекция этого вектора в пространство любого неприводимого представления $T^{(\lambda)}(g)$, входящего в состав $T(g)$, будет инвариантным вектором соответствующего (приводимого) представления $T^{(\lambda)}(k)$ группы $\mathcal{K}$. Рассуждая далее так же как при выводе формулы (2.26) мы найдем, что произвольная положительно определенная функция на X, удовлетворяющая (3.I), здесь будет представима формулой

(3.I5) $\mathcal{B}(x_1,x_2)=\int_{\bar{\mathcal{G}}_{\mathcal{K}}} Tr\left(\mathcal{P}_{\mathcal{K}}^{(\lambda)}T^{(\lambda)}(g_2^{-1}g_1)\mathcal{P}_{\mathcal{K}}^{(\lambda)}\mathcal{F}_{\mathcal{K}}(d\lambda)\right),$

где $\bar{\mathcal{G}}_{\mathcal{K}}$ – подмножество тех $\lambda\in\bar{\mathcal{G}}$, для которых в пространстве $H^{(\lambda)}$ имеется хотя бы один вектор, инвариантный относительно всех преобразований $T^{(\lambda)}(k),\ k\in\mathcal{K}$; $\mathcal{P}_{\mathcal{K}}^{(\lambda)}$ – оператор проектирования в $H^{(\lambda)}$ на максимальное инвариантное относительно всех $T^{(\lambda)}(k)$ подпространство $H_{\mathcal{K}}^{(\lambda)}$, $\mathcal{F}_{\mathcal{K}}(d\lambda)$ – эрмитова неотрицательно определенная "операторная мера" на $\bar{\mathcal{G}}_{\mathcal{K}}$ со значениями из кольца операторов в пространстве $\mathcal{H}_{\mathcal{K}}^{(\lambda)}$, а g_1 и g_2 - произвольные элементы из классов смежности по $\mathcal{K}$, задающих точки x_1 и x_2 из X. Оператор $\mathcal{P}_{\mathcal{K}}^{(\lambda)} T^{(\lambda)}(g) \mathcal{P}_{\mathcal{K}}^{(\lambda)}$ в формуле (3.15) естественно рассматривается как оператор в подпространстве $\mathcal{H}_{\mathcal{K}}^{(\lambda)}$; интеграл по $\bar{\mathcal{Y}}_{\mathcal{K}}$ здесь следует понимать как сумму интегралов, распространенных по подмножествам $\bar{\mathcal{Y}}_n \subset \bar{\mathcal{Y}}_{\mathcal{K}}$, $n = \infty, 1, 2, \ldots$ таким, что при $\lambda \in \bar{\mathcal{Y}}_n$ подпространство $H_{\mathcal{K}}^{(\lambda)}$ является n-мерным.

Таким образом, имеет место следующая

Теорема 6. Если группа движений $\mathcal{G}$ является сепарабельной локально компактной группой типа I, то класс инвариантных относительно всех движений положительно определенных функций на однородном пространстве $X = \mathcal{G}/\mathcal{K}$, совпадает с классом функций, представимых в виде (3.15), где $\mathcal{F}(d\lambda)$ - эрмитовская неотрицательно определенная "операторная мера" на $\bar{\mathcal{Y}}_{\mathcal{K}}$ (со значениями, являющимися операторами в $H_{\mathcal{K}}^{(\lambda)}$) такая, что

$$\text{(3.16)} \qquad \operatorname{Tr} \mathcal{F}(\bar{\mathcal{Y}}_{\mathcal{K}}) < \infty .$$

Выбрав далее в каждом из подпространств $H_{\mathcal{K}}^{(\lambda)}$, отвечающих всевозможным $\bar{\mathcal{Y}}_n$, определенный базис мы можем переписать формулу (3.15) в виде

$$\text{(3.17)} \qquad \mathcal{B}(x_1, x_2) = \int_{\bar{\mathcal{Y}}_{\mathcal{K}}} \sum_{ij} \Phi_{ij}^{(\lambda)}(x_1, x_2) \mathcal{F}_{ji}(d\lambda),$$

непосредственно обобщающем (3.9). Здесь через $\{\Phi_{ij}^{(\lambda)}(x_1, x_2)\}$ обозначено полное семейство линейно независимых зональных сферических функций, отвечающих неприводимому представлению $T^{(\lambda)}(g)$; это семейство будет конечным (n^2-членным) при $\lambda \in \bar{\mathcal{Y}}_n$, $n = 1, 2, \ldots$, и бесконеным при $\lambda \in \bar{\mathcal{Y}}_\infty$.

Существует еще один важный класс однородных пространств, для которых можно выписать общую формулу (и притом даже более простую чем (3.15)) для произвольных положительно определенных функций

$\mathcal{B}(x_1, x_2)$, удовлетворяющих (3.I). Этим классом являются симметрические однородные пространства Э.Картана. Однородное пространство $X = \mathcal{G}/\mathcal{K}$ называется симметрическим, если в группе $\mathcal{G}$ его движений имеется инволютивный автоморфизм $g \to g'$ (т.е.изоморфное отображение $\mathcal{G}$ на $\mathcal{G}$, при котором $(g')' = g$), выделяющий стационарную подгруппу $\mathcal{K}$, т.е.такой, что $g' = g$ тогда и только тогда, когда $g \in \mathcal{K}$. Нетрудно видеть, что этому условию будут удовлетворять, например, любые однородные пространства постоянной кривизны (см.,например, [33] , где можно найти также ряд других примеров однородных симметрических пространств). Для произвольных однородных симметрических пространств Гельфандом [34] (см.также книгу [I8] , § 3I) была доказана следующая важная теорема,существенно дополняющая теорему 6:

Теорема 6'. В случае симметрического однородного пространства $X = \mathcal{G}/\mathcal{K}$ каждому неприводимому унитарному представлению $T^{(\lambda)}(g)$ группы $\mathcal{G}$ отвечает не более чем одна зональная сферическая функция $\Phi^{(\lambda)}(x_1, x_2)$,т.е.подпространство $H_{\mathcal{K}}^{(\lambda)}$ при любом $\lambda \in \bar{\mathcal{G}}$ здесь не более чем одномерно. Любая инвариантная относительно движений положительно определенная функция $\mathcal{B}(x_1, x_2)$ на таком пространстве X может быть представлена формулой

$$(3.I8) \quad \mathcal{B}(x_1, x_2) = \int_{\bar{\mathcal{G}}_1} \Phi^{(\lambda)}(x_1, x_2) \mathcal{F}(d\lambda),$$

где $\mathcal{F}(d\lambda)$ - неотрицательная мера на $\bar{\mathcal{G}}_{\mathcal{K}}$ (совпадающем в этом случае с $\bar{\mathcal{G}}_{\mathcal{K}}$) такая,что интеграл в правой части (3.I7) сходится.

В случае симметрического пространства X функции $\Phi^{(\lambda)}(x_1, x_2)$ будут симметрично зависеть от x_1 и x_2 ,так как здесь всегда существует движение, меняющее порядок этих двух точек, иначе говоря, они будут зависеть лишь от сложного расстояния между точками x_1 и x_2 (см. [33]).

Перейдем теперь к "спектральному разложению" самого однородного случайного поля $\xi(x)$. Предположим, что корреляционная функция $\mathcal{B}(x_1, x_2)$ этого поля может быть представлена в виде (3.I5) (частным случаем этой формулы является и (3.I8)). Будем рассуждать аналогично доказательству теоремы . Рассмотрим гильбертово пространство $L^2(\mathcal{F}_{\mathcal{X}})$ операторных функций $\mathcal{U}^{(\lambda)}$, $\lambda \in \bar{\mathcal{Y}}_{\mathcal{X}}$, со значениями из кольца ограниченных операторов, действующих из $H_{\mathcal{X}}^{(\lambda)}$ в $H^{(\lambda)}$, и нормой

$$(3.19)\quad \|\mathcal{U}^{(\lambda)}\|^2 = \int_{\bar{\mathcal{Y}}_{\mathcal{X}}} \operatorname{Tr}\left[\mathcal{U}^{(\lambda)} \mathcal{F}_{\mathcal{X}}(d\lambda)\, \mathcal{U}^{(\lambda)*}\right].$$

Любой ограниченный оператор V в $H^{(\lambda)}$ мы можем, если угодно, рассматривать как "оператор из $H_{\mathcal{X}}^{(\lambda)}$ в $H^{(\lambda)}$", ограничив его область определения подпространством $H_{\mathcal{X}}^{(\lambda)}$; во избежание путаницы этот "суженный" оператор мы будем обозначать через $V\mathcal{P}_{\mathcal{X}}^{(\lambda)}$. В таком случае соответствие

$$(3.20)\quad T^{(\lambda)}(g)\mathcal{P}_{\mathcal{X}}^{(\lambda)} \longleftrightarrow \xi(x) \qquad \text{при } x = g'\mathcal{K} \ni g$$

будет изометрическим отображением множества $\{\xi(x), x \in X\}$ пространства $\mathfrak{h}$ в $L^2(\mathcal{F}_{\mathcal{X}})$, которое можно продолжить до изометрического отображения $L^2(\mathcal{F}_{\mathcal{X}})$ в $\mathfrak{h}$. Если теперь паре векторов $f_1 \in H^{(\lambda)}$, $f_2 \in H_{\mathcal{X}}^{(\lambda)}$ и измеримому множеству $\Lambda \in \bar{\mathcal{Y}}_{\mathcal{X}}$ мы сопоставим $\mathcal{U}^{(\lambda)}(\Lambda; f_1, f_2) \in L^2(\mathcal{F}_{\mathcal{X}})$ с помощью формулы (2.30), то при отображении (3.20)

$$(3.21)\quad \mathcal{U}^{(\lambda)}(\Lambda; f_1, f_2) \longleftrightarrow Z(\Lambda; f_1, f_2) = (Z(\Lambda) f_1, f_2),$$

где $Z(\Lambda)$ – случайный линейный оператор из $H^{(\lambda)}$ в $H_{\mathcal{X}}^{(\lambda)}$, зависящий от $\Lambda \in \bar{\mathcal{Y}}_{\mathcal{X}}$. Далее так же как при выводе формулы (2.32) доказывается, что

$$(3.22)\quad \xi(x) = \int_{\bar{\mathcal{Y}}_{\mathcal{X}}} \operatorname{Tr}\left(Z(d\lambda)\, T^{(\lambda)}(g)\mathcal{P}_{\mathcal{X}}^{(\lambda)}\right), \qquad x = \{g'\mathcal{K}\} \ni g$$

и что при любых $f_1,\ g_1 \in H^{(\lambda)},\quad f_2,\ g_2 \in H_{\mathcal{X}}^{(\lambda)}$

(3.23) $E(Z(\Lambda_1)f_1,f_2)\overline{(Z(\Lambda_2)g_1,g_2)}=(f_1,g_1)(\mathcal{F}_{\mathcal{K}}(\Lambda_1\cap\Lambda_2)g_2,f_2).$

Отсюда легко получается следующая

Теорема 7. Для того, чтобы случайное поле $\xi(x)$ на однородном пространстве $X=\mathcal{G}/\mathcal{K}$, группа $\mathcal{G}$ движений которого является сепарабельной локально компактной группой типа I, было однородным необходимо и достаточно, чтобы его можно было представить в виде (3.22), где $Z(\Lambda)$ - случайный линейный оператор из $H^{(\lambda)}$ в $H^{(\lambda)}_{\mathcal{K}}$, аддитивно зависящий от множества $\Lambda\subset\bar{\mathcal{G}}_{\mathcal{K}}$ и удовлетворяющий при любых $f_1,g_1\in H^{(\lambda)}$, $f_2,g_2\in H^{(\lambda)}_{\mathcal{K}}$ условию (3.23). Здесь $\mathcal{F}_{\mathcal{K}}(\Lambda)$- эрмитов неотрицательный оператор в $H^{(\lambda)}_{\mathcal{K}}$ удовлетворяющий (3.16), через который корреляционная функция $\mathcal{B}(x_1,x_2)$ поля $\xi(x)$ выражается при помощи формулы (3.15).

Аналогичные результаты справедливы и для однородных случайных полей на произвольном симметрическом однородном пространстве X, с той только разницей, что в этом случае пространство $H^{(\lambda)}_{\mathcal{K}}$ одномерно при любом $\lambda\in\bar{\mathcal{G}}_{\mathcal{K}}=\bar{\mathcal{G}}_1$ и, следовательно, $Z(\Lambda)$- случайный линейный функционал в $H^{(\lambda)}$, удовлетворяющий условию

(3.24) $EZ(\Lambda_1)f_1\overline{Z(\Lambda_2)g_1}=\mathcal{F}(\Lambda_1\cap\Lambda_2)(f_1,g_1),$

где $\mathcal{F}(\Lambda)$- неотрицательная мера в $\bar{\mathcal{G}}_1$, фигурирующая в формуле (3.18).

В "координатной записи" формулы (3.22) и (3.23) будут, разумеется, весьма похожи на (3.7) и (3.8). В частности, в случае симметрического пространства X эти формулы обращаются в

(3.25) $$\xi(x)=\int_{\bar{\mathcal{G}}_1}\sum_n\Phi_n^{(\lambda)}(x)Z_n(d\lambda),$$

(3.26) $$EZ_n(\Lambda_1)\overline{Z_m(\Lambda_2)}=\delta_{nm}\mathcal{F}(\Lambda_1\cap\Lambda_2),$$

где $\{\Phi_n^{(\lambda)}(x),\ n=1,2,\ldots\}$ - полная система "сферических функций", отвечающих зональной сферической функции $\Phi^{(\lambda)}(x_1,x_2)$. Теорему 7 (вместе с теоремами 6 и 6') можно рассматривать как обобщение приведенных выше теорем I, 3, 4 и 5.

Примеры. Известно, что в n-мерном евклидовом пространстве R_n (с обычной группой движений $\mathcal{G}$) имеется однопараметрическое семейство зональных сферических функций $\Phi^{(\lambda)}(r)$ (зависящих лишь от расстояния r между точками x_1 и x_2):

$$(3.27)\quad \Phi^{(\lambda)}(r) = J_{\frac{n-2}{2}}(\lambda r) \Big/ (\lambda r)^{\frac{n-2}{2}}\ ;\quad 0 \le \lambda < \infty,$$

(см., например, [35], [28]). Таким образом, формула Шенберга (I.4) является частным случаем общей формулы (3.18).

Общие сферические функции на евклидовой плоскости R_2 выписаны в работе Крейна [35] (см. также Виленкин [24]). Подставив эти формулы в (3.25) - (3.26) мы получим следующее общее представление однородных и изотропных случайных полей на плоскости:

$$(3.28)\quad \xi(r,\varphi) = \sum_{n=-\infty}^{\infty} e^{-in\varphi} \int_0^{\infty} J_n(r\lambda) Z_n(d\lambda),$$

где (r,φ) - полярные координаты на плоскости и $Z_n(d\lambda)$ удовлетворяет (3.26). Аналогично этому для однородных и изотропных полей в n-мерном евклидовом пространстве R_n исходя из результатов заметки [28] (см. также [32], том П) может быть получено представление:

$$(3.29)\quad \xi(r,\theta_1,\ldots,\theta_{n-2},\varphi) =$$

$$= \Sigma A_{\ell,m_1,\ldots,m_{n-3},\pm m_{n-2}} \mathcal{Y}_{\ell,m_1,\ldots,m_{n-3},\pm m_{n-2}}(\theta_1,\ldots,\theta_{n-2},\varphi) \int \frac{J_{\ell+\frac{n-2}{2}}(r\lambda)}{(r\lambda)^{\frac{n-2}{2}}} Z_{\ell,m_1,\ldots,m_{n-3},\pm m_{n-2}}(d\lambda),$$

где $r,\theta_1,\ldots,\theta_{n-2},\varphi$ - сферические координаты в R_n, суммирование в правой части распространено по всем $\ell=0,1,2,\ldots, 0\le m_{n-2}\le m_{n-3}\le\ldots$ $\ldots\le m_1\le\ell$ и двум знакам m_{n-2}, $A_{\ell,m_1,\ldots,m_{n-3},\pm m_{n-2}}$ - нормированные константы, просто выражающиеся через Γ-функцию, $\mathcal{Y}_{\ell,m_1,\ldots,m_{n-3},\pm m_{n-2}}$ - соответствующие поверхностные гармоники, а $Z_{\ell,m_1,\ldots,m_{n-3},\pm m_{n-2}}(d\lambda)$ - счетное семейство взаимно некоррелированных случайных мер на прямой $[0,\infty]$ с одинаковым математическим ожиданием квадрата модуля.

В случае n-мерного пространства Лобачевского L_n зональные сферические функции имеют вид

$$(3.30)\quad \Phi^{(\lambda)}(r)=\frac{\mathcal{P}_{\mu(\lambda)}^{-\frac{n-2}{2}}(ch r)}{(ch r)^{\frac{n-2}{2}}},\qquad \mu(\lambda)=-\frac{1}{2}+i\sqrt{\lambda-\frac{n^2}{4}+\frac{n}{2}-\frac{1}{4}}\ ,$$

где $\mathcal{P}_\mu^\nu$ – специальное решение дифференциального уравнения Лежандра (см.Крейн [35] , Виленкин [29] ; для трехмерного пространства Лобачевского, где $\Phi^{(\lambda)}(r)=\frac{\sin\sqrt{\lambda-1}\,r}{\sqrt{\lambda-1}\,sh r}$, соответствующий результат еще раньше был получен Гельфандом и Наймарком [36], [18]). Отсюда для корреляционной функции изотропного поля L_n получается следующая общая формула (впервые указанная Крейном [35]) :

$$(3.31)\quad B(r)=\int_0^\infty \frac{\mathcal{P}_{\mu(\lambda)}^{-\frac{n-2}{2}}(ch r)}{(ch r)^{\frac{n-2}{2}}}\,d\Phi(\lambda),\qquad \mu(r)=-\frac{1}{2}+i\sqrt{\lambda-\frac{n^2}{4}+\frac{n}{2}-\frac{1}{4}}\ .$$

В работе [35] выписаны также все незональные сферические функции пространства L_2 ; подставляя их в (3.25) мы получим следующий общий вид однородного случайного поля на плоскости Лобачевского:

$$(3.32)\quad \xi(r,\varphi)=\sum_{n=-\infty}^{\infty}\gamma_n e^{-in\varphi}\int_0^\infty \mathcal{P}^n_{-\frac{1}{2}+\sqrt{\frac{1}{4}-\lambda}}(ch r)\,Z_n(d\lambda),\qquad \gamma_n=\sqrt{(-1)^n\frac{\Gamma(\sqrt{\frac{1}{4}-\lambda}-m+\frac{1}{2})}{\Gamma(\sqrt{\frac{1}{4}-\lambda}+m-\frac{1}{2})}}\ ,$$

где $Z_n(d\lambda)$ удовлетворяют (3.26). Для n-мерного пространства L_n все зональные функции указаны в заметке [29] ; отсюда для произвольного изотропного поля в L_n получается представление типа (3.29), в котором только функцию $\frac{J_{\ell+\frac{n-2}{2}}(r\lambda)}{(r\lambda)^{\frac{n-2}{2}}}$ следует заменить на $\frac{\mathcal{P}_{\mu(\lambda)}^{-\frac{n-2}{2}-\ell}(ch r)}{(sh r)^{\frac{n-2}{2}}}$ (и изменить значение констант $A_{\ell,m_1,\dots,m_{n-3},\pm m_{n-2}}$).

Ряд других примеров полных систем зональных сферических функций на частных однородных пространствах указан в работах [26], [37] , [38] ; в [33] изучены также некоторые общие свойства таких функций, существенно облегчающие их нахождение. Вопрос о нахождении произвольных (незональных) сферических функций является более сложным; однако, для конкретных пространств $X = \mathcal{G}/\mathcal{K}$ он также в ряде случаев может быть решен вполне эффективно.

4. Многомерные однородные поля.

Дальнейшие обобщения.

Понятие однородного случайного поля, рассмотренное в разделах 2 и 3, допускает ряд дальнейших обобщений. Сейчас мы вкратце остановимся на некоторых из них.

(a) Многомерные однородные случайные поля на группах. Пусть $\xi(g) = \{\xi_1(g), \dots, \xi_N(g)\}$ – N-мерное случайное поле на группе $\mathcal{G}$, а $\{\mathcal{U}(g)\}$ – какое-то (не обязательно унитарное) N-мерное представление этой группы. Поле $\xi(g)$ называется однородным относительно левых сдвигов полем величин ξ, преобразующихся по представлению $\mathcal{U}(g)$, если для всех $g, g_1, g_2 \in \mathcal{G}$

(4.1) $E\xi_i(g_1) = E[V_g \xi(g_1)]_i$, $E\xi_i(g_1)\overline{\xi_j(g_2)} = E[V_g\xi(g_1)]_i\overline{[V_g\xi(g_2)]_j}$,

где

(4.2) $$V_g\xi(g_1) = \mathcal{U}(g)\xi(g^{-1}g_1).$$

Иначе говоря, если $M(g) = E\xi(g)$ – вектор средних значений поля $\xi(g)$, а $\mathcal{B}(g_1, g_2) = \|E\xi_i(g_1)\overline{\xi_j(g_2)}\|$ – его корреляционная матрица, то

(4.3) $$M(g_1) = \mathcal{U}(g)M(g^{-1}g_1) = \mathcal{U}(g_1)M^{(0)},$$

(4.4) $$\mathcal{B}(g_1, g_2) = \mathcal{U}(g)\mathcal{B}(g^{-1}g_1, g^{-1}g_2)\mathcal{U}^*(g) = \mathcal{U}(g_2)\mathcal{B}^{(0)}(g_2^{-1}g_1)\mathcal{U}^*(g_2),$$

где $M^{(0)}$ – некоторый постоянный вектор, а $\mathcal{B}^{(0)}(g)$ – матрица, зависящая от одного оргумента. Аналогично определяется и поле $\xi(g)$, однородное относительно правых сдвигов: в этом случае надо только положить: $V_g\xi(g_1) = \mathcal{U}(g)\xi(g_1 g)$.

Следуя Колмогорову (см.Розанов [39]) многомерное поле $\xi(g)$ можно также интерпретировать как заданное на $\mathcal{G}$ поле линейных операторов $\xi_g(a)$ из некоторого линейного пространства $\mathcal{A}$ в пространство случайных величин $h_{\mathcal{Y}}$:

$$(4.5) \quad \xi_g(a) = \sum_k \xi_k(g) a_k, \qquad a = (a_1, a_2, \dots) \in \mathcal{A}.$$

В этом случае поле $\xi_g(a)$ будет называться однородным относительно левых сдвигов, если существует представление $\{U^+(g)\}$ группы $\mathcal{G}$ в пространстве $\mathcal{A}$ такое, что для всех $g, g_1, g_2 \in \mathcal{G}$

$$(4.6) \quad E\xi_{g_1}(a) = E\xi_{gg_1}(U^+(g)a), \quad E\xi_{g_1}(a_1)\overline{\xi_{g_2}(a_2)} = E\xi_{gg_1}(U^+(g)a_1)\overline{\xi_{gg_2}(U^+(g)a_2)}$$

(представление $U^+(g)$ связано с $U(g)$ соотношением: $U^+(g) = [U(g^{-1})]'$, т.е. $U^+_{ij}(g) = U_{ji}(g^{-1})$).

Такая "бескоординатная" интерпретация многомерных полей особенно удобна при изучении полей бесконечномерных.

Общий вид вектора средних значений однородного поля величин ξ непосредственно определяется формулой (4.3): вектор $M^{(0)}$ в этой формуле может быть любым постоянным вектором из $\mathcal{A}$. Остается еще только выяснить общий вид матрицы $\mathcal{B}(g_1, g_2)$ (или, что то же самое, матрицы $\mathcal{B}^{(0)}(g)$). С этой целью мы дополнительно предположим, что группа $\mathcal{G}$ - сепарабельная локально компактная группа типа I (в частности, она может быть произвольной компактной группой). Далее пусть a - произвольный постоянный вектор из $\mathcal{A}$ и $a(g) = U^+(g)a$. Тогда в силу (4.6) случайное поле

$$(4.7) \quad \xi_a(g) = \sum_k \xi_k(g) a_k(g)$$

будет однородным относительно левых сдвигов одномерным полем, линейно зависящим от параметров a_k . Применив к этому полю теорему 4, получим

$$(4.8) \quad \xi_a(g) = \int_{\hat{\mathcal{G}}} \mathrm{Tr}\{T^{(\lambda)}(g) Z_a(d\lambda)\},$$

где

(4.9) $Z_a(\Lambda) = \sum_k^m Z_k(\Lambda) a_k,$

а $\mathbf{Z}(\Lambda) = \{Z_1(\Lambda), \ldots, Z_N(\Lambda)\}$ – векторная "случайная операторная мера" на $\mathcal{G}$ такая, что

(4.10) $E\left(Z_m(\Lambda_1) f_1, f_2\right)\overline{\left(Z_n(\Lambda_2) g_1, g_2\right)} = (f_1, g_1)\left(\mathcal{F}_{mn}(\Lambda_1 \cap \Lambda_2) g_2, f_2\right),$

т.е., если $Z_m(\Lambda) = \| Z_{ij,m}(\Lambda) \|,$

(4.11) $E Z_{ij,m}(\Lambda_1) \overline{Z_{k\ell,n}(\Lambda_2)} = \delta_{j\ell} \mathcal{F}_{ik,mn}(\Lambda_1 \cap \Lambda_2),$

(4.12) $E Z_a(\Lambda_1) f_1, f_2)\overline{(Z_b(\Lambda_2) g_1, g_2)} = (f_1, g_1)(\mathcal{F}(\Lambda_1, \Lambda_2)(a, g_2), (b, f_2)).$

Здесь $\mathcal{F}(\Lambda) = \| \mathcal{F}_{ik,mn}(\Lambda) \|$ – это аддитивно зависящий от Λ эрмитовский неотрицательный оператор с конечным следом в "кронекеровском произведении" $A \times H^{(\lambda)}$ пространств A и $H^{(\lambda)}$ (т.е. в прямой сумме изоморфных $H^{(\lambda)}$ пространств $e_1 H^{(\lambda)}, e_2 H^{(\lambda)}, \ldots$, где $\{e_1, e_2, \ldots\}$ – базис в A). Формулу (4.8) можно также переписать в виде

(4.13) $\xi(g) = \int_{\bar{\mathcal{G}}} \mathcal{U}(g) \operatorname{Tr}\left(T^{(\lambda)}(g) \mathbf{Z}(d\lambda)\right)$

или же

(4.14) $\xi_m(g) = \sum_n \mathcal{U}_{mn}(g) \int_{\bar{\mathcal{G}}} \sum_{i,j} T_{ij}^{(\lambda)}(g) Z_{ji,n}(d\lambda).$

Равенство (4.3) при этом эквивалентно условию

(4.15) $E\mathbf{Z}(\Lambda) = \begin{cases} M^{(0)}, & \text{если } \lambda_0 \subset \Lambda, \\ 0, & \text{если } \lambda_0 \bar{\subset} \Lambda, \end{cases}$

где $\{T^{(\lambda_0)}(g)\}$ –единичное представление группы $\mathcal{G}$, так что $T^{(\lambda_0)}(g) = I$. Для корреляционной матрицы $\mathcal{B}(g_1, g_2)$ из (4.8) и (4.12) вытекает соотношение

(4.16) $\mathcal{B}(g_1, g_2) = \mathcal{U}(g_1) \int_{\bar{\mathcal{G}}} \operatorname{Tr}\{T^{(\lambda)}(g_2^{-1} g_1) \mathcal{F}(d\lambda)\} \mathcal{U}^*(g_2) =$

$$= \mathcal{U}(g_2) \cdot \mathcal{U}(g_2^{-1} g_1) \int \operatorname{Tr}\{T^{(\lambda)}(g_2^{-1} g_1) \mathcal{F}(d\lambda)\} \cdot \mathcal{U}^*(g_2)$$

или иначе

$$(4.17)\ \mathcal{B}_{mn}(g_1,g_2)=\sum_{s,t}U_{ms}(g_1)\overline{U_{nt}(g_2)}\int_{\bar{\mathcal{G}}}\sum_{i,j}T_{ij}^{(\lambda)}(g_2^{-1}g_1)\mathcal{F}_{ji,st}(d\lambda),$$

$$(4.18)\ \mathcal{B}_{mn}^{(0)}(g)=\sum_{s}U_{ms}(g)\int_{\bar{\mathcal{G}}}\sum_{i,j}T_{ij}^{(\lambda)}(g)\mathcal{F}_{ji,sn}(d\lambda).$$

Для того,чтобы это соотношение было совместимо с (4.3), должно только выполняться неравенство:

$$(4.19)\qquad \mathcal{F}_{mn}(\lambda_0)\geqslant M_m^{(0)}\overline{M_n^{(0)}}.$$

Отсюда легко выводится

Теорема 8. Для того,чтобы многомерное случайное поле $\xi(g)$ на ~~группе~~ сепарабельной локально компактной группе $\mathcal{G}$ типа I было однородным относительно левых сдвигов полем величин ξ ,преобразующихся по представлению $\{U(g)\}$, необходимо и достаточно,чтобы оно было представимо в виде (4.13), где $Z(\Lambda)$ - векторная "случайная операторная мера" на $\bar{\mathcal{G}}$, удовлетворяющая (4.12), а $\mathcal{F}(\Lambda)$ - "операторная мера" на $\bar{\mathcal{G}}$, значения которой являются эрмитовскими неотрицательными операторами с конечным следом в "кронекеровском произведении" $A\times H^{(\lambda)}$ пространств A и $H^{(\lambda)}$. Среднее значение $M=E\xi(g)$ поля $\xi(g)$ при этом дается формулой (4.3), где $M^{(0)}$ определяется из (4.15), а корреляционная матрица $\mathcal{B}(g_1,g_2)$ - формулой (4.16).

Обратно, любая матрица вида (4.16) является корреляционной матрицей некоторого многомерного однородного случайного поля; среднее значение этого поля может принимать любое значение вида (4.3), где $M^{(0)}$ - постоянный вектор, удовлетворяющий (4.19).

(в) Многомерные однородные поля на однородных пространствах. Многомерное поле $\xi(x)=\{\xi_1(x),\xi_2(x),\dots\}$ на $X=\mathcal{G}/\mathcal{X}$ называется однородным, если его первые и вторые моменты не меняются при применении к величинам $\xi(x)$ преобразования $\xi(x)\to U(g)\xi(g^{-1}x)$, где $\{\mathcal{U}(g)\}$ - некоторое представление группы $\mathcal{G}$. Ясно, что если поле $\xi(x)$ будет однородным, то вектор $M(x) = E\xi(x)$ и матрица $\mathcal{B}(x_1, x_2) = \|E\xi_i(x_1)\overline{\xi_j(x_2)}\|$ будут удовлетворять соотношениям:

(4.20) $M(x) = \mathcal{U}(g)M(g^{-1}x)$,

(4.21) $\mathcal{B}(x_1, x_2) = \mathcal{U}(g)\mathcal{B}(g^{-1}x_1, g^{-1}x_2)\mathcal{U}^*(g)$

(ср.(4.3)-(4.4)). Иначе это еще можно выразить так: линейный оператор из $\mathcal{A}$ в $\mathfrak{h}_y$

(4.22) $\xi_x(a) = \sum_k \xi_k(x) a_k$

будет в этом случае обладать свойствами:

(4.23) $E\xi_x(a) = E\xi_{gx}(\mathcal{U}^+(g)a)$, $E\xi_{x_1}(a_1)\overline{\xi_{x_2}(a_2)} = E\xi_{gx_1}(\mathcal{U}^+(g)a_1)\overline{\xi_{gx_2}(\mathcal{U}^+(g)x_2)}$.

Из формулы (4.20) сразу следует, что

(4.24) $M(x) = \mathcal{U}(g)M^{(0)}$,

где $M^{(0)}$ - произвольный вектор из $\mathcal{A}$, инвариантный относительно всех преобразований $\mathcal{U}(k)$, $k \in \mathcal{K}$:

(4.25) $\mathcal{U}(k)M^{(0)} = M^{(0)}$ при всех $k \in \mathcal{K}$.

Значительно более сложным является вопрос об общем виде корреляционной матрицы $\mathcal{B}(x_1, x_2)$ и связанный с ним вопрос о "спектральном разложении" самого поля $\xi(x)$. Мы разберем сейчас этот вопрос для случая конечномерного поля $\xi(x)$ на компактном однородном пространстве X; после этого уже нетрудно будет сообразить, как должны писаться соответствующие формулы и в более общем локально компактном случае.

Итак, пусть $\xi(x) = \{\xi_1(x), \dots, \xi_N(x)\}$, где $x \in \mathcal{G}/\mathcal{K}$, $\mathcal{G}$ - компактная группа. Поле $\xi(x)$ можно также рассматривать как N-мерное однородное случайное поле $\xi(g)$ на $\mathcal{G}$, постоянное на всех левых классах смежности $\mathcal{G}$ по $\mathcal{K}$. Следовательно (см.(4.I4),(4.II))

(4.26) $\eta_i(g)=\sum_j \xi_j(g)\mathcal{U}^+_{ji}(g)=\sum_j \mathcal{U}_{ij}(g^{-1})\xi_j(g)$

при любом $i=1,2,\ldots,N$ будет допускать представление

(4.27) $\eta_i(g)=\sum_\lambda \sum_{m,n} z^{(\lambda)}_{nm,i}\, T^{(\lambda)}_{mn}(g)$,

где $\{T^{(\lambda)}(g)\}$ – всевозможные неприводимые (и неэквивалентные) унитарные представления $\mathcal{G}$ и

(4.28) $E z^{(\lambda)}_{nm,i}\, \overline{z^{(\mu)}_{sr,j}} = \delta_{\lambda\mu}\delta_{mr} f^{(\lambda)}_{ns,ij}$.

Но в нашем случае $\xi_k(g)=\xi_k(gk)$, $k\in\mathcal{K}$ поэтому при любом $k\in\mathcal{K}$

(4.29) $\eta_i(gk)=\sum_j \xi_j(gk)\mathcal{U}^+_{ji}(gk)=\sum_j\sum_\ell \xi_j(g)\mathcal{U}^+_{j\ell}(g)\mathcal{U}^+_{\ell i}(k)=\sum_\ell \eta_\ell(g)\mathcal{U}^+_{\ell i}(k)$,

т.е.

(4.30) $\sum_\lambda\sum_{m,n}\sum_r z^{(\lambda)}_{nm,i}\, T^{(\lambda)}_{mr}(g) T_{rn}(k)=\sum_\lambda\sum_{m,n}\sum_\ell z^{(\lambda)}_{nm,\ell}\, T^{(\lambda)}_{mn}(g)\mathcal{U}^+_{\ell i}(k)$.

Обратно, из (4.30) следует, что $\xi_j(gk)\equiv\xi_j(g)$;таким образом,условие (4.30) является необходимым и достаточным для того,чтобы поле $\xi(g)$ в (4.26) – (4.27) было на самом деле однородным многомерным полем на $X=\mathcal{G}/\mathcal{K}$.

В силу соотношений ортогональности (2.8) из (4.30) вытекает, что при любых λ, m и n для всех $k\in\mathcal{K}$

(4.31) $\sum_r z^{(\lambda)}_{nm,i}\, T^{(\lambda)}_{nr}(k)=\sum_\ell z^{(\lambda)}_{nm,\ell}\,\mathcal{U}^+_{\ell i}(k)$.

Выберем теперь и в пространстве A , в котором действует представление $\mathcal{U}^+(g)$, и в пространствах представлений $T^{(\lambda)}(g)$, $\lambda=1,2,\ldots$, базисы таким образом,чтобы эти представления распались на неприводимые представления подгруппы $\mathcal{K}$. Помимо того мы потребуем еще чтобы эквивалентные между собой представления $\mathcal{K}$, входящие в $\mathcal{U}^+$ и $T^{(\lambda)}$ записывались одинаково – этого также всегда можно добиться простым изменением базисов. Пусть представление $\mathcal{U}^+(g)$ группы $\mathcal{G}$ распадается на неприводимые представления $\{V^{(i)}(k)\}$, $i=1,\ldots,J$;

кратность представления $V^{(i)}(k)$ в $\mathcal{U}^{+}(g)$ мы обозначим через $\mathcal{L}_i$, а его размерность – через S_i. В таком случае индекс i ($=1,\ldots,\mathcal{N}$) при компонентах вектора ξ или η удобно будет заменить составным индексом i_{ls} ($i=1,\ldots,\mathcal{J}$; $l=1,\ldots,\mathcal{L}_i$; $s=1,\ldots,S_i$); $\mathcal{U}^{+}_{ij}(k)$ при этом обратится в

(4.32) $\mathcal{U}^{+}_{i_{ls},j_{mt}}(k)=\delta_{ij}\delta_{lm}V^{(i)}_{st}(k)$.

Аналогично этому если представление $T^{(\lambda)}(g)$ группы $\mathcal{G}$ распадается на неприводимые представления $V^{(n)}(k)$, $n=1,\ldots,\mathcal{N}_\lambda$,причем кратность $V^{(n)}(k)$ и $T^{(\lambda)}(g)$ равна $u_{n\lambda}$, то вместо индекса n ($=1,\ldots,d_\lambda$), нумерующего компоненты матриц $T^{(\lambda)}(g)$, мы будем употреблять составной индекс n_{ua} ($n=1,\ldots,\mathcal{N}_\lambda$; $u=1,\ldots,u_{n\lambda}$; $a=1,\ldots,S_a$); тогда

(4.33) $T^{(\lambda)}_{n_{ua},m_{wb}}(k)=\delta_{nm}\delta_{uw}V^{(n)}_{ab}(k)$.

Подставляя (4.32) и (4.33) в (4.31), без труда получаем

(4.34) $z_{n_{ua},m,i_{ls}}=\delta_{ni}\delta_{as}z^{(\lambda,i)}_{m,ul}$,

где

(4.35) $E z^{(\lambda,i)}_{m,ul}\overline{z^{(\mu,j)}_{n,wk}}=\delta_{\lambda\mu}\delta_{mn}f^{(\lambda;ij)}_{uw,lk}$.

Отсюда следует,что

(4.36) $\eta_{i_{ls}}(g)=\sum_{\lambda}\sum_{m=1}^{d_\lambda}\sum_{u=1}^{u_{i\lambda}}z^{(\lambda,i)}_{m,ul}T^{(\lambda)}_{mi_{us}}(g)$,

т.е. каждая компонента $\eta_j(g)$ вектора $\eta(g)$ разлагается лишь по тем столбцам матриц $T^{(\lambda)}(g)$, которые принадлежат тому же неприводимому представлению группы $\mathcal{K}$, что и j-й столбец матрицы $\mathcal{U}^{+}(g)$ и занимают в нем те же место, что и этот j-й столбец. В силу (4.36) компоненты $\xi_j(x)$ поля $\xi(x)$ будут иметь вид

(4.37) $\xi_j(x)=\sum z^{(\lambda,i)}_{m,ul}\,\mathcal{U}_{ji_{ls}}(g)T^{(\lambda)}_{mi_{us}}(g)$, где сумма берется по всем дважды повторяющимся индексам. Нетрудно проверить,что функции $U_{ji_{es}}(g)T^{(\lambda)}_{mi_{us}}(g)$ на самом деле будут зависеть лишь от класса смежности $\mathcal{G}$ по $\mathcal{K}$, т.е. фактически будут функциями от $x \in X$, а не от $g \in \mathcal{G}$. Из (4.37) и (4.35) немедленно получается и общая формула для корреляционной матрицы $\mathcal{B}(x_1, x_2) = \| \mathcal{B}_{jk}(x_1, x_2) \|$:

$$(4.38) \quad \mathcal{B}_{jk}(x_1, x_2) = \sum f^{(\lambda, il)}_{uv, ab} U_{jias}(g_1) T^{(\lambda)}_{lvtius}(g_2^{-1} g_1) U^*_{lvtk}(g_2),$$

где суммирование также ведется по всем дважды повторяющимся индексам и $\| f^{(\lambda, il)}_{uv, ab} \|$ – неотрицательно определенная матрица по всем своим индексам:

$$(4.39) \quad \sum_{i,l} \sum_{u,v} \sum_{a,b} f^{(\lambda; il)}_{av, ab} \alpha_{iua} \overline{\alpha_{lvb}} \geqslant 0.$$

Равенство (4.37) и задает общее "спектральное представление" многомерного однородного поля на компактном однородном пространстве X, а равенство (4.38) – "спектральное разложение" соответствующей корреляционной матрицы.

В случае сепарабельной локально компактной группы $\mathcal{G}$ типа I и ее компактной подгруппы $\mathcal{K}$ спектральное разложение многомерного однородного поля $\xi(x)$ на $X = \mathcal{G}/\mathcal{K}$ и его корреляционной матрицы $\mathcal{B}(x_1, x_2)$ будет даваться формулами, полностью аналогичными (4.37), (4.38) и (4.39); только суммирование по λ здесь придется заменить интегрированием соответствующих функций множества $d\lambda \subset \overline{\mathcal{G}}$ по всему пространству $\overline{\mathcal{G}}$.

В применении к конкретным многообразиям X и типам величин (т.е.представлениям $U(g)$) общие формулы могут,естественно, существенно упроститься. Так,например,в случае полей на сфере S_2 трехмерного эвклидова пространства R_3 стационарная подгруппа $\mathcal{K} = O_2$ (подгруппа вращений вокруг оси) оказывается коммутативной, так что все ее неприводимые представления одномерны (имеют вид $V^{(m)}(\varphi)=e^{im\varphi}$, где φ - угол поворота). Кроме того в этом случае во все неприводимые представления группы $\mathcal{G}=O_3$ полной группы вращений R_3) каждое представление $V^{(m)}$ группы O_2 входит не более, чем по разу. Поэтому здесь вместо составных индексов $i_{\varrho s}$, $n_{u\alpha}$ и т.д. обычно можно ограничиться простыми индексами i, n и т.д. Воспользовавшись далее формулами для всех матричных элементов $T^{(\lambda)}_{mn}$ представлений группы O_3 , указанными в [23] , нетрудно написать явный вид соотношений (4.37) и (4.38) для случая, например, векторных однородных случайных полей на S_2 или тензорных однородных полей невысокого ранга. Те же результаты можно получить, воспользовавшись описанным в [23] инвариантным относительно вращений разложением векторных и тензорных (не случайных) полей на сфере S_2 по соответственно подобранным специальным функциям (см.также заметку [40] , в которой разобрана задача о подобном же инвариантном разложении векторных полей на сфере S_n $(n+1)$-мерного пространства R_n).

В другом важном частном случае - случае векторных однородных и изотропных полей в n-мерном эвклидовом пространстве $R_n = M_n/O_n$ (M_n - группа движений R_n , O_n - группа n-мерных вращений) - задача упрощается тем,что векторное представление $\mathcal{U}$ группы M_n является в то же время неприводимым унитарным представлением подгруппы O_n , не более раза входящим в каждое неприводимое представление M_n . Поэтому формула (4.36) здесь принимает вид

$$(4.40)\qquad \eta_s(g) = \int_{\bar{\mathcal{G}}} \sum_m T^{(\lambda)}_{ms}(g) Z_m(d\lambda)$$

(интеграл берется по всем неприводимым представлениям M_n, содержащим векторное представление подгруппы O_n). Воспользовавшись далее результатами заметки [28] отсюда можно получить как общее спектральное представление корреляционной матрицы $\mathcal{B}_{ij}(x_1, x_2)$ векторного поля (найденное иным способом в [4]), так и спектральное представление самого поля $\xi(x)$.

(c) Обобщенные однородные случайные поля. Если пространство $X = \mathcal{G}/\mathcal{K}$ (где $\mathcal{K}$ может быть и единичной подгруппой) является конечномерным дифференцируемым многообразием (группа $\mathcal{G}$ является группой Ли), то наряду с обычными случайными полями $\xi(x)$ (или $\xi(x) = \{\xi_1(x), \dots, \xi_n(x)\}$) на X можно рассматривать также и обобщенные случайные поля (случайные распределения) в смысле Ито [41] – Гельфанда [42]. Согласно [41], [42] обобщенным полем называется случайный линейный функционал $\xi(\varphi)$ (или $\xi(\varphi) = \{\xi_1(\varphi), \dots, \xi_n(\varphi)\}$) относительно функции $\varphi \in \mathcal{D}$, где $\mathcal{D}$ – введенный Л.Шварцем класс бесконечно дифференцируемых комплексных функций точки x, обращающихся в нуль вне некоторого компакта. Обобщенное поле $\xi(\varphi)$ называется однородным, если математические ожидания $m(\varphi) = E\xi(\varphi)$ и $\mathcal{B}(\varphi_1, \varphi_2) = E\xi(\varphi_1)\overline{\xi(\varphi_2)}$ таковы, что при любом $\varphi \in \mathcal{G}$

(4.41) $m(\varphi) = m(V_g\varphi), \quad \mathcal{B}(\varphi_1, \varphi_2) = \mathcal{B}(V_g\varphi_1, V_g\varphi_2),$

где $V_g\varphi(x) = \varphi(g^{-1}x)$. В многомерном случае эти условия заменяются условиями

(4.42) $M(\varphi) = U(g)M(V_g\varphi), \quad \mathcal{B}(\varphi_1, \varphi_2) = U(g)\mathcal{B}(V_g\varphi_1, V_g\varphi_2)U^*(g),$

где $M(\varphi)$ – вектор средних значений поля $\xi(\varphi)$, $\mathcal{B}(\varphi_1, \varphi_2)$ – его корреляционная матрица, а $U(g)$ – некоторое представление группы $\mathcal{G}$. Почти все результаты настоящей статьи сохраняются и для таких обобщенных однородных полей, если только полученные выше формулы для $\xi(x)$ (или для $\xi(g)$) понимать в том смысле, что

$$(4.43) \qquad \xi(\varphi) = \int_X \xi(x)\varphi(x)\,dx,$$

где dx инвариантная относительно преобразований $g \in \mathcal{G}$ мера в X. Единственным отличием будет то, что при таком подходе к нашим формулам выражения для $\xi(x)$ (или $\xi(g)$) вполне могут быть и расходящимися – лишь бы только они становились сходящимися после интегрирования по $\varphi(x)dx$ или $\varphi(g)dg$. Поэтому условия сходимости (2.13), (2.27),(3.16) и т.п.для обобщенных полей уже не будут обязательными и должны будут быть заменены на менее жесткие условия. Точный вид этих ослабленных условий будет определяться асимптотическими свойствами соответствующих матричных элементов $T_{ij}^{(\lambda)}(g)$ и сферических функций $\Phi_{ij}^{(\lambda)}(x)$; в наиболее важных конкретных случаях эти асимптотические свойства обычно известны или могут быть без труда получены, так что изучение соответствующих однородных обобщенных полей здесь не связано ни с какими новыми трудностями (ср.,например, работы [5] и [4], посвященные обобщенным случайным полям в евклидовых пространствах R_n).

Новые постановки вопросов возникают, если наряду с обобщенными полями $\xi(\varphi)$, заданными на функциональном пространстве $\mathfrak{D}$ бесконечно дифференцируемых функций рассматривать также и обобщенные поля на каких-либо иных функциональных пространствах – например, на пространстве $\mathfrak{D}_n$ функций, дифференцируемых не более n раз (и обращающихся в нуль вне компакта) или на классах функций, введенных в работах Гельфанда и Шилова (см. [43]). Ясно, что сужению класса функций φ будет при этом отвечать расширение соответствующего класса однородных случайных полей,т.е.ослабление условий регулярности, заменяющих (2.27), (3.16) и т.д.Класс обыкновенных (не обобщенных) однородных случайных полей является, с этой точки зрения, пересечением целого семейства охватывающих друг друга классов обобщенных случайных полей, отвечающих все более и более суживающимся классам функций φ. Непосредственно за классом обыкновенных случайных полей в этом семействе будет следовать класс однородных случайных мер - случайных функций $\xi(S)$ множества $S \subset X$ таких, что $E\xi(S)=E\xi(gS)$, $E\xi(S_1)\overline{\xi(S_2)} = E\xi(gS_1)\overline{\xi(g_1S_2)}$ при любом $g \in \mathcal{G}$ (подчеркнем, во избежание недоразумения, что термин "случайная мера" здесь имеет другой смысл, чем раньше в этой статье, где он применялся лишь к случайным функциям множеств, принимающим на непересекающихся множествах некоррелированные значения). Класс случайных мер можно рассматривать как класс обобщенных случайных полей, заданных на всевозможных непрерывных функциях φ; поэтому в отличие от последующих классов обобщенных полей он может быть определен на любом топологическом однородном пространстве (т.е. группа $\mathcal{G}$ в этом случае вовсе не должна обязательно являться группой Ли).

В том же смысле, что и для обобщенных однородных полей Ито-Гельфанда над пространством $\mathcal{D}$, основные результаты настоящей статьи сохраняются и для всех остальных классов обобщенных полей. Однако вопрос о точном виде соответствующих "условий ограниченности" (налагаемых на коэффициенты $f_{ij}^{(\lambda)}$ или на "операторную меру" $\mathcal{F}(d\lambda)$) для каждого из этих классов должен решаться особо.

(d) Поля со случайными однородными приращениями. В теории случайных процессов наряду с процессами стационарными хорошо изучены также и более общие процессы, имеющие стационарные приращения (см., например, [44], [8]). Аналогичное обобщение может быть предложено и в отношении понятия однородного случайного поля на произвольном однородном пространстве X. А именно, поле $\xi(x)$ мы будем называть полем со случайными однородными приращениями, если всевозможные разности $\xi(x_1)-\xi(x_2)=\xi(x_1,x_2)$ будут представлять собой однородное (относительно преобразований $g(x_1,x_2)=(gx_1,gx_2)$) случайное поле на пространстве $X\times X$.

Не следует считать, что теория полей с однородными приращениями непосредственно сводится к теории однородных полей на новом однородном пространстве: это не так, ибо в пространстве $X\times X$ группа $\mathcal{G}=\{g\}$ уже не будет транзитивной. Поэтому в общем случае установление "спектральных разложений" для полей с однородными приращениями требует привлечения некоторых новых соображений.

Основными численными характеристиками поля с однородными приращениями будут первые и вторые моменты разностей $\xi(x_1,x_2)$:

(4.44) $m(x_1,x_2)=E\xi(x_1,x_2),\ \mathcal{B}(x_1,x_2;x_3,x_4)=E\xi(x_1,x_2)\overline{\xi(x_3,x_4)}$.

Функция $m(x_1,x_2)$ полностью характеризуются тем, что она удовлетворяет функциональному уравнению $m(x_1,x_2)+m(x_2,x_3)=m(x_1,x_2)$ и условию $m(gx_1,gx_2)=m(x_1,x_2)$; отсюда обычно уже не трудно определить общий вид функций. Что же касается вторых моментов $\mathcal{B}(x_1,x_2;x_3,x_4)$, то все они без труда выражаются через свои значения $\mathcal{B}(x_1,x_2;x_3,x_2)$ при $x_4=x_2$; если поле $\xi(x)$ — вещественное, то в силу алгебраического тождества

$$(a-b)(c-d)=\frac{1}{2}\left[(a-d)^2+(b-c)^2-(a-c)^2-(b-d)^2\right]$$

мы можем даже ограничиться лишь изучением функций

(4.45) $\mathcal{B}(x_1,x_2)=E\left[\xi(x_1)-\xi(x_2)\right]^2$,

зависящих всего от двух переменных. Как было выяснено Шенбергом [45] функции $\mathcal{B}(x_1,x_2)$ полностью характеризуются следующим свойством (родственным свойству положительной определенности): при любых $n, x_1, \ldots, x_n \in \mathcal{G}$ и любых вещественных числах $a_1, \ldots, a_n$ таких, что $\sum_i a_i = 0$, должно выполняться неравенство

$$(4.46) \qquad \sum_{i,k=1}^{n} \mathcal{B}(x_i, x_k) a_i a_k \leq 0.$$

Таким образом описание всех (вещественных) полей с однородными приращениями в известном смысле эквивалентно описанию всех функций $\mathcal{B}(x_1, x_2)$, удовлетворяющих (3.I) и (4.46).

В случае компактного пространства X нетрудно показать, что класс таких функций $\mathcal{B}(x_1, x_2)$ совпадает с классом функций вида $\mathcal{B}_1(x, x) - \mathcal{B}_1(x_1, x_2)$, где x - произвольная точка X, а $\mathcal{B}_1(x_1, x_2)$ - инвариантная (в смысле (3.I)) положительно определенная функция на X (см. Бохнер [II]). Отсюда сразу следует, что в компактном случае класс полей с однородными приращениями совпадает с классоом просто однородных полей.

Для более общих локально компактных пространств X последнее утверждение уже оказывается неверным: здесь класс полей с однородными приращениями может оказаться существенно шире класса однородных полей. Так, в частности, обстоит дело в случае евклидова пространства $X = R_n$ (с группой трансляций или общей группой движений в качестве группы $\mathcal{G}$); см. по этому поводу работу [4] Поля с однородными приращениями в пространстве R_n (одномерные и векторные многомерные) играют существенную роль в статистической теории турбулентности; именно в этой связи впервые было сформулировано и само определение таких полей (Колмогоров [46]). В еще более частном случае $X = R_1$ (т.е. для случайных процессов $\xi(t)$) изучен также и более общий класс процессов с однородными приращениями произвольного порядка ([4I], [47], [48]); теория таких процессов также может быть перенесена и на случай полей на некоторых отличных от R_1 однородных пространствах.

В случае, когда группа $\mathcal{G}$ является группой Ли, удобно рассматривать сразу обобщенные случайные поля с однородными приращениями. Нетрудно понять, что такие поля можно определить,как поля $\xi(\varphi)$, удовлетворяющие соотношениям (4.4I)(или (4.42)), но заданные лишь на подпространстве $\mathcal{D}_1$ функций из $\mathcal{D}$ таких,что

$$\int \varphi(x)\,dx = 0 \tag{4.47}$$

(ср. [4]). Это определение допускает дальнейшие обобщения, связанные с рассмотрением однородных случайных полей на некоторых других линейных подпространствах пространства $\mathcal{D}$; в частном случае полей на прямой R_1 на этом пути также можно построить теорию полей с однородными приращениями n-го порядка.

ЛИТЕРАТУРА

I. I.J.Schoenberg,"Metric spaces and completely monotone functions",Annals of Math.,vol.39 (1938),pp.811-841.

2. А.М.Яглом, "Однородная и изотропная турбулентность в вязкой сжимаемой жидкости", Известия Акад. Наук СССР, сер. географ. и геофиз., т.I2 /I948/, стр. 50I-522.

3. J.E.Moyal, "The spectra of turbulence in a compressible fluid; eddy turbulence and random noise",Proc.Cambr. Philos.Soc.,vol.48 (1952),pp.329-344.

4. А.М.Яглом, "Некоторые классы случайных полей в n-мерном пространстве, родственные стационарным случайным процессам", Теория вероятн. и ее применения, т.2 /I957/, стр. 292-338.

5. K.Itô, "Isotropic random current",Proc.3-d Berkeley Symposium on Math.Stat and Probab.,Berkeley and Los Angeles,vol.2(1956),pp.125-132.

6. А.М.Обухов, "Статистически однородные случайные поля на сфере", Успехи матем. наук, т.2, № 2 /I947/, стр. I96-I98.

7. I.J.Schoenberg, "Positive definite functions on spheres", Duke Math.Journal,vol.9(1942),pp.96-108.

8. J.L.Doob, "Stochastic processes",New York,1953.

9. L.Pontrjagin,"Topological groups",Princeton,1939.

I0. A.Weil, "L'integration dans les groupes topologiques et ses applications", Paris, 1940.

II. S.Bochner, "Hibbert distances and positive definite functions", Annals of Math., vol.42(1941), pp.647-656.

I2. K.Karhunen, "Ueber lineare Methoden in der Wahrscheinlichkeitsrechnung", Annals Acad. Sci.Fennicae, A, No.37 (1947), pp.3-79.

I3. H.Cramer, "A contribution to the theory of stochastic processes", Proc.2-nd Berkeley Symp.on Mathem. Stat.and Probab., Berkeley and Los Angeles, 1951, pp.329-339.

I4. Д.А.Райков, "Гармонический анализ на коммутативных группах с мерой Хаара и теория характеров", Труды математ. института имени В.А.Стеклова, № I4 /I945/, 86 стр.

I5. J.Kampé de Ferriet, "Analyse harmonique des fonctions aléatoires stationnaires d'ordre 2 sur un groupe abélien localement compact", C.R.Acad.Sci., Paris, vol.226 (1948), pp.868-870.

I6. A.Blanc-Lapierre, R.Fortet, "Théorie des fonctions aléatoires", Paris, 1953.

I7. И.М.Гельфанд, Д.А.Райков, "Неприводимые унитарные представления локально бикомпактных групп", Матем.сборник, т. I3 /55/ /I943/, стр. 30I-3I6.

I8. М.А. Наймарк, "Нормированные кольца", Москва, 1956,

I9. Harish-Chandra, "Representations of a semisimple Lie group on a Banach space", Trans.Amer.Math.Soc., vol.75 (1953), pp.185-243.

20. J.Dixmier, " Sur les représentations unitaires des groupes de Lie algébriques",Ann.Inst.Fourier, vol.7 (1957),pp.315-328.

21. G.W.Mackey,"Borel structure in groups and their duals", Trans.Amer.Math.Soc.,vol.85 (1957),pp.134-165.

22. A.Guichardet, "Sur une probléme posé par G.W.Mackey", C.R.Acad.Sci.,Paris,vol.250 (1960),pp.962-963.

23. И.М.Гельфанд, Р.А.Минлос, З.Я.Шапиро, "Представления группы вращений и группы Лоренца",Москва,1958.

24. Н.Я.Виленкин, "Бесселевы функции и представления группы евклидовых движений", Успехи матем.наук, т.II /1956/, стр. 69-112.

25. V.Bargman," Irreducible unitary representations of the Lorentz group",Annals of Math.,vol48 (1947), pp.568-640.

26. И.М.Гельфанд, М.А.Наймарк, "Унитарные представления классических групп," Труды математ. института имени В.А.Стеклова, № 36, /1950/, 288 стр.

27. Н.Я.Виленкин, Э.Л.Аким, А.А.Левин "Матричные элементы неприводимых унитарных представлений группы евклидовых движений трехмерного пространства и их свойства", Доклады Акад. наук СССР, т. 112 /1957/, стр. 987-989.

28. Н.Я.Виленкин, "Матричные элементы неприводимых унитарных представлений группы вещественных ортогональных матриц и группы движений $(n-1)$ -мерного евклидова пространства",Доклады Акад.наук СССР,т.113 /1957/, стр. 16-19.

29. Н.Я.Виленкин, "Матричные элементы неприводимых унитарных представлений группы движений пространства Лобачевского и обобщенные преобразования Фока-Мелера" Доклады Акад. наук СССР, т. II8/I958/, стр. 2I9-222.

30. E.Cartan, "Sur la détermination d'un systéme orthogonal complet dans un espace de Riemann symmétrique clos", Rend.Circ.Mat.Palermo, vol53 (1929), pp.217-252.

3I. H.Weyl, "Harmonics on homogeneous manifolds", Annals of Math., vol.35 (1934), pp.486-494.

32. A.Erdélyi, "Higher transcendental functions", vol.I-III, New York, 1953-1955.

33. Ф.А.Березин, И.М.Гельфанд, "Несколько замечаний к теории сферических функций на симметрических римановых многообразиях", Труды Московского Матем.Общ-ва, т. 5 /I956/, 3I2-35I.

34. И.М.Гельфанд, "Сферические функции на симметрических римановых пространствах", "Доклады Акад. наук СССР, т. 70 /I950/, стр. 5-8.

35. М.Г.Крейн, "Эрмитово-положительные ядра на однородных пространствах", ч. I-П, Украинский матем.журнал, т. I, № 4 /I949/, стр. 64-98; т. 2 № I/I950/, стр. I0-59.

36. И.М.Гельфанд, М.А.Наймарк, "Унитарные представления группы Лоренца", Известия Акад.наук СССР, сер. матем., т. II/I947/, стр. 4II-504.

37. Ф.А.Березин, Ф.И.Карпелевич, "Зональные сферические функции и операторы Лапласа на некоторых симметрических пространствах", Доклады Акад.наук СССР, т.II8 /I958/, стр. 9-I2.

38. А.Н.Богаевский, "Вычисление зональных сферических функций" Доклады Акад.наук СССР, т. I29/I959/, стр. 484-487.

39. Ю.А.Розанов, "Спектральная теория многомерных стационарных процессов с дискретным временем", Успехи матем. наук, т.I3, /I958/, стр. 93-I42.

40. А.А.Кириллов, "Представления группы вращений n-мерного евклидова пространства сферическими векторами полями", Доклады Акад.наук СССР, т. II6/I957/, стр.538-54I.

4I. K.Itô, "Stationary random distributions", Mem.College Sci. Univ.Kyoto, Ser.A vol.28 (1953), pp.209-223.

42. И.М.Гельфанд, "Обобщенные случайные процессы", Доклады Акад. наук СССР, т. I00/I955/, стр. 853-856.

43. И.М.Гельфанд, Г.Е.Шилов, "Пространства основных и обобщенных функций" /"Обобщенные функции", вып.2/, Москва I958.

44. А.Н.Колмогоров, "Кривые в гильбертовском пространстве, инвариантные по отношению к однопараметрической группе движений", Доклады Акад.наук СССР, т.26 /I940/, стр. 6-9.

45. I.J.Schoenberg, "Metric spaces and positive definite functions",Trans.Amer.Math.Soc.,vol.44 (1938), pp.522-536.

46. А.Н.Колмогоров, "Локальная структура турбулентности в несжимаемой жидкости при очень больших числах Рейнольдса", Доклады Акад. наук СССР, т. 30 /I94I/, стр. 299-303.

47. А.М.Яглом, "Корреляционная теория процессов со случайными стационарными n-ми приращениями", Математ. сборник, т. 37 /79/, /I955/, стр. I4I-I96.

48. М.С. Пинскер, "Теория кривых в гильбертовом пространстве со стационарными n-ми приращениями",Известия Акад. Наук СССР, сер. матем.т.I9/I955/,3I9-345.

ON THE PROBABILITY OF LARGE DEVIATIONS FOR THE SUMS OF INDEPENDENT VARIABLES

Yu. V. Linnik

USSR Academy of Sciences

1. Introduction

The classical theory of the summation of independent random variables as expounded in the book [8] in its simplest case considers the increasing sums $S_n = X_1 + \cdots + X_n$. For the properly normed and centered sums $Z_n = S_n/B_n - A_n$ the behavior in the limit of the probability measures generated by $\{Z_n\}$ on the real axis is studied.

The most general theorems are the integral theorems on the limit behavior of

(1.1) $$P\{Z_n < x\}.$$

Although the theory of local limit theorems is rather well developed [8], it is not yet of such finished character as that of integral limit theorems. Limit theorems for the expression (1.1) usually suppose that $n \to \infty$ and x is a fixed number.

However, many problems occurring in such different fields as mathematical statistics ([4], [2]), information theory ([5], [19]), statistical physics of polymers [18], rubber chemistry [17], and even analytical arithmetic require certain information on the limit behavior of (1.1) not contained in the classical limit theorems. The information required concerns the asymptotic behavior of

$$P\{Z_n > x\} \tag{1.2}$$

for "large values" of x, that is for $x = x_n$ increasing as n increases; the corresponding problems will be called problems on of the probability of large deviations. As probabilities of events of this kind are generally small, ~~in general~~, the usual methods of establishing the limit theorems (characteristic functions, partial differential equations) are too rough to give satisfactorily general results and the desired asymptotic results were considered in the literature under certain very stringent conditions imposed upon the variables X_j.

The first theorem on ~~the~~ the probability of large deviations was published by A.I. Khintchine [10] in 1929 and related to the particular case of Bernoulli variables. The same case was treated more completely by N.V. Smirnov [16]. In 1938 appeared the fundamental paper [4] of H. Cramér containing the first result of a general nature in the theory of large ~~big~~ deviations. It was improved by W. Feller [7] and by V.V. Petrov [12].

Now, we shall formulate Petrov's result, restricting ourselves to the case of identically distributed variables for the sake of simplicity. Let X_1, X_2, $\cdots$, X_n be a sequence of independent identically distributed variables with

$$
(1.3)\quad
\begin{cases}
E(X_j) = 0, & D(X_j) = \sigma^2, \\
S_n = X_1 + \cdots + X_n, & Z_n = \dfrac{S_n}{\sigma\sqrt{n}}, \\
F_n(x) = P\{Z_n < x\}, & G(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{x} e^{-u^2/2}\,du.
\end{cases}
$$

Cramér's [11] condition (C)

$$(1.4)\qquad E \exp(a|X_j|) < \infty$$

must hold for some $a > 0$. Then for $x \geq 1$, $x = o(\sqrt{n})$, and $n \to \infty$ we have

$$(1.5)\qquad \frac{1 - F_n(x)}{1 - G(x)} = \exp\left[\frac{x^3}{\sqrt{n}}\lambda\left(\frac{x}{\sqrt{n}}\right)\right]\left[1 + O\left(\frac{x}{\sqrt{n}}\right)\right],$$

$$(1.6)\qquad \frac{F_n(-x)}{G(x)} = \exp\left[-\frac{x^3}{\sqrt{n}}\lambda\left(\frac{x}{\sqrt{n}}\right)\right]\left[1 + O\left(\frac{x}{\sqrt{n}}\right)\right],$$

where $\lambda(z)$ is a power series involving the cumulants of the variables X_j and convergent for $|z| \leq \varepsilon_0$, $\varepsilon_0 > 0$. Later on, W. Richter [13], [14], [15] introduced systematically the saddlepoint method into the theory of large deviations. For a particular case, this was done earlier by H. Daniels [4]. Under Cramér's condition (C), Richter deduced several local limit theorems for large ~~the big~~ deviations, established the connection of Cramér's method with the saddlepoint method, and investigated the necessity of condition (C) for the formulas (1.5) and (1.6) to hold for $x = o(\sqrt{n})$.

All the results hitherto obtained used Cramér's condition (C). The analytical meaning of the condition (C) is that the characteristic function (ch.f.) of the X_j is analytical in some neighborhood of ~~the~~ zero, and so in the corresponding strip. This enables us to apply complex function theory and the saddlepoint method.

But if the condition (C) is violated, the methods hitherto applied fail. The purpose of this paper is to give some applications of a new approach which enables us to obtain rather general results. Of the class of problems subject to this method we shall treat here only the problem of ~~the~~ normal convergence and the problem of limit theorems valid for all values of x for $n \to \infty$.

2. Zone of normal convergence: integral limit theorems

We consider here the normal convergence problem for large ~~big~~ deviations for the sake of simplicity only for independent identically distributed variables.

Let $X_1, X_2, \cdots, X_n, \cdots$ be independent identically distributed variables with $E(X_j) = 0$, $D(X_j) = 1$, $Z_n = (X_1 + X_2 + \cdots + X_n)/\sqrt{n}$. Let $\Psi(n) \to \infty$ be any monotone function. The sequence of the segments $[0, \Psi(n)]$ will be called ~~the~~ a zone of normal convergence (z.n.c.) if, for $n \to \infty$,

$$(2.1) \qquad \frac{P\{Z_n > x\}}{\frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-u^2/2}\, du} \to 1$$

for any $x \in [0, \Psi(n)]$. ~~The~~ A z.n.c. $[-\Psi(n), 0]$ is defined similarly. The definition does not require the convergence to be uniform, although in all the theorems obtained it will be.

Under Cramér's condition (C) for any $\Psi(n) = o(n^{1/6})$, both $[0, \Psi(n)]$ and $[-\Psi(n), 0]$ will be z.n.c. The zones with $\Psi(n) = o(n^{1/6})$ will be called narrow zones. Our two first theorems relate to the zones with $\Psi(n) = n^\alpha$, where $\alpha > 0$ is a constant.

Theorem 1. If for any $\alpha < 1/2$, the zone $[0, n^\alpha]$ and the zone $[-n^\alpha, 0]$ are z.n.c., then all the variables X_j are normal.

Of course, this is also sufficient for the zones $[0, n^\alpha]$ and $[-n^\alpha, 0]$ to be z.n.c. This last fact is trivial. Thus, we see that it is sufficient to investigate the values of $\alpha < 1/2$.

Theorem 2. Let $\rho(n) \to \infty$ be a monotone function increasing as slowly as we please and let $0 < \alpha < 1/2$. If

$\alpha < 1/6$, ~~the~~ a necessary condition for both the ~~the~~ zones $[0, n^{\alpha}\rho(n)]$ and $[-n^{\alpha}\rho(n), 0]$ to ~~both~~ be z.n.c. is

$$(2.2) \qquad E \exp |X_j|^{\frac{4\alpha}{2\alpha+1}} < \infty .$$

This condition is sufficient for ~~the~~ the zones $[0, n^{\alpha}/\rho(n)]$ and $[-n^{\alpha}/\rho(n), 0]$ to be z.n.c. and the convergence is then uniform. If $1/6 \leq \alpha < 1/2$, consider the sequence of the critical numbers

$$(2.3) \qquad \frac{1}{6}, \frac{1}{4}, \frac{3}{10}, \ldots, \frac{1}{2}\,\frac{s+1}{s+3} \ldots \longrightarrow \frac{1}{2}.$$

Let

$$(2.4) \qquad \frac{1}{2}\,\frac{s+1}{s+3} \leq \alpha < \frac{1}{2}\,\frac{s+2}{s+4}.$$

If the zones $[0, n^{\alpha}\rho(n)]$ and $[-n^{\alpha}\rho(n), 0]$ are z.n.c., the condition (2.2) must hold and moreover all the moments of X_j up to $(s + 3)$# must coincide with the moments of the normal law. These two conditions are sufficient for the zones $[0, n^{\alpha}/\rho(n)]$ and $[-n^{\alpha}/\rho(n), 0]$ to be z.n.c. This convergence is then uniform.

As the normal law is completely determined by the sequence of the corresponding moments, theorem 1 is an immediate consequence of theorem 2.

We consider now the narrow zones with $\Psi(n) = o(n^{1/6})$ other than $[0, n^{\alpha}]$. ~~The~~ A condition necessary for the zones $[0, \Psi(n)\rho(n)]$ and $[-\Psi(n)\rho(n), 0]$ and sufficient for the zones $[0, \Psi(n)/\rho(n)]$ and $[-\Psi(n)/\rho(n), 0]$ to be z.n.c. is of the type

$$E \exp h(|X_j|) < \infty, \tag{2.5}$$

where $h(x)$ is a monotone function depending upon $\Psi(n)$.

It is simpler, however, to describe $\Psi(n)$ in terms of $h(x)$. To this ~~effect~~ end we consider several classes of functions of $h(x)$. The functions $h(x)$ will be assumed to be positive, monotone, and differentiable.

Class $\underline{\mathrm{I}}$ will denote the functions $h(x)$ ~~under~~ satisfying the condition

$$(\cancel{\ln}\,\log x)^{2+\zeta_0} \leq h(x) \leq x^{1/2}, \qquad x \geq 1. \tag{2.6}$$

Here $\zeta_0 > 0$ is any small fixed number. Functions increasing faster than $x^{1/2}$ will not be required for the narrow zone investigations.

Class $\underline{\mathrm{II}}$ consists of the functions $h(x)$ under the condition

$$\rho_1(x)\,\cancel{\ln}\,\log x \leq h(x) \leq (\cancel{\ln}\,\log x)^{2+\zeta_0}, \tag{2.7}$$

where $x \geq 1$, $\rho_1(x)$, $\rho_2(x)$, $\cdots$ in what follows are given positive monotone functions increasing as slowly as we please.

Class $\underline{\mathrm{III}}$ consists of functions $h(x)$ such that

$$3\,\cancel{\ln}\,\log x \leq h(x) \leq M\,\cancel{\ln}\,\log x, \tag{2.8}$$

where $M > 3$ is a given constant. The inequality $h(x) \geq 3\,\cancel{\ln}\,\log x$ is connected with the existence of the third moment.

Consider now the functions of Class $\underline{\mathrm{I}}$ as defined by (2.6). We put

(2.9) $$h(x) = \exp\,[H(\cancel{\ln}\ \log x)].$$

Then $H(z)$ is a monotone differentiable function. We introduce the following supplementary conditions

(2.10) $$H'(z) \leq 1,$$

(2.11) $$H'(z)\exp\left[H(z)\right] \to \infty \quad z \to \infty.$$

These conditions follow from (2.6) if $H'(z)$ is assumed to be monotone; otherwise we adopt them to simplify the results.

Given a function $h(x)$ of class I, under the supplementary conditions (2.10) and (2.11), or of class II, we determine new functions $\Lambda(n)$ by means of the equation

(2.12) $$h[\sqrt{n}\,\Lambda(n)] = [\Lambda(n)]^2.$$

Theorem 3. The condition

(2.13) $$E\exp h(|X_j|) < \infty,$$

where $h(x)$ belongs to class I [with (2.10) and (2.11)] or class II, is necessary for the zones $[0, \Lambda(n)\rho(n)]$ and $[-\Lambda(n)\rho(n), 0]$ to be z.n.c. and sufficient for the zones $[0, \Lambda(n)/\rho(n)]$ and $[-\Lambda(n)/\rho(n), 0]$ to be z.n.c. The convergence in this case is uniform.

We pass now to the functions $h(x)$ belonging to class III. This case can be studied by classical means [8]. For the sake of completeness we formulate

Theorem 4. Condition (2.13), where $h(x)$ belongs to class III, is necessary for the zones $\cancel{[0, \sqrt{\ln n}\,\rho(n)]}$ and

$[0, \rho(n)(\log n)^{1/2}]$

$[-\rho(n)(\log n)^{1/2}, 0]$

$\cancel{[-\sqrt{\ln n}\,\rho(n), 0]}$ to be z.n.c. and sufficient for the zones $[-(\log n)^{1/2}/\rho(n), 0]$ $[0, \cancel{\sqrt{\ln n}}\,(\log n)^{1/2}/\rho(n)]$ and $\cancel{[-\sqrt{\ln n}/\rho(n), 0]}$ to be z.n.c.

Thus, if $E|X_j|^M = \infty$ for a fixed M, the z.n.c. cannot be essentially wider than $[0, \cancel{\sqrt{\ln n}}\,(\log n)^{1/2}]$. It is, roughly speaking, of this size if $E|X_j|^M < \infty$ with $M \geqq 3$. The case $E|X_j|^3 = \infty$ (nonexistence of the third moment) was studied by several authors (see [8] for the literature).

The case (2.8), which is class III of the functions h(x), corresponds to slowly decreasing "probability tails" $P\{X_j > x\}$. In this case, as we shall show later, a new type of limit theorems holds: limit theorems valid for the whole x-axis.

3. Local limit theorems

In the preceding section we considered integral limit theorems. We pass now to the local limit theorems relating to normal convergence. These theorems are usually considered for variables possessing a probability density or for ~~integral~~ integer valued variables. We can consider also ~~the~~ probability measures on the ring of ~~the integral numbers~~ integers of an algebraic number field. We shall restrict ourselves to the class (d) of all random variables possessing a continuous bounded density g(x). Then Z_n (see section 2) will also have a continuous density $p_{Z_n}(x)$. The zone $[0, \Psi(n)]$ will be called ~~the~~ a zone of uniform local normal convergence (z.u.l.n.c.) if

$$(3.1) \qquad \frac{p_{\widehat{Z_n}}(x)}{\frac{1}{\sqrt{2\pi}} e^{-x^2/2}} \to 1 ,$$

as $n \to \infty$, uniformly for $x \in [0, \Psi(n)]$. The z.u.l.n.c. $[-\Psi(n), 0]$ are defined similarly.

Theorem 5. For the variables X_j belonging to the class (d) the z.u.l.n.c. behave with respect to the necessary and sufficient conditions indicated in theorems 1 to 4 in the same way as the z.n.c. for the general random variables in theorems 1 to 4.

~~In fact,~~ The local limit theorems for ~~big~~ large deviations are easier to prove than the corresponding integral ones, by the method proposed here. ~~Namely~~ In fact, the existence of the probability density $g(x)$ greatly facilitates the proof.

BF

4. ~~Uniform local normal convergence~~ Proof of necessity of condition (2.2)

We shall be able to expound here the proofs for only the simplest cases so as to present the new approach in its most transparent form; the proofs of all the theorems 1 to 5, although not basically different, are more involved and will be published elsewhere. In particular, we shall treat only the zones $[0, n^\alpha]$ and $[-n^\alpha, 0]$ with $0 < \alpha < 1/2$ and only local limit theorems. However, since a part of the necessary conditions for the integral theorems is almost trivial, we shall begin by dwelling upon it.

~~In MS [and ρ(n)], "0" not minus~~

Let the zones $[0, n^{\alpha}\rho(n)]$ and $[-n^{\alpha}\rho(n), 0]$ be z.n.c. We shall prove that

$$(4.1) \qquad E \exp\left[|X_j|^{\frac{4\alpha}{2\alpha+1}}\right] < \infty.$$

Suppose (4.1) does not hold. Then it is easy to see that there exists either ~~the~~ a sequence $x_m \to \infty$ such that

$$(4.2) \qquad P\{X_1 > x_m\} > \exp\left[-2x_m^{\frac{4\alpha}{2\alpha+1}}\right]$$

or ~~the~~ a sequence $-x_m \to -\infty$ such that

$$(4.3) \qquad P\{X_1 < -x_m\} > \exp\left[-2x_m^{\frac{4\alpha}{2\alpha+1}}\right]$$

Suppose that (4.2) holds. For a sufficiently large m, choose n such that $x_m = n^{\alpha+1/2}\rho(n)$. The zone $[0, n^{\alpha}\rho(n)]$ being ~~the~~ a z.n.c., we must have

$$(4.4) \qquad P\left\{Z_n > \frac{n^{\alpha}\rho(n)}{2}\right\} \cancel{< \exp}\left\{-\frac{n^{2\alpha}[\rho(n)]^2}{16}\right\}.$$

But the event $Z_n > n^{\alpha}\rho(n)/2$ will surely occur if the two independent events $X_1 > n^{\alpha+1/2}\rho(n)$ and $|(X_2 + X_3 + \cdots + X_n)/\sqrt{n}| < 1$ occur simultaneously. Hence, by the central limit theorem, where $c_0 > 0$ is a constant, Corrected acc. to MS. OK? No!

$$(4.5) \qquad P\left\{Z_n > \frac{n^{\alpha}\rho(n)}{2}\right\} > c_0\, P\left\{X_1 > n^{\alpha+1/2}\rho(n)\right\}$$

$$> c_0 \exp -2n^{2\alpha} [\rho(n)]^{\frac{4\alpha}{2\alpha+1}}$$

by (4.2). ~~As~~ Since $\alpha < 1/2$, ~~thus~~ it follows that $4\alpha/(2\alpha+1) < 1$ and (4.5) contradicts (4.4). The case (4.3) is treated similarly. The proof of the necessity of (4.1) for the zones $[0, n^{\alpha}\rho(n)]$ and $[-n^{\alpha}\rho(n), 0]$ to be z.u.l.n.c. is constructed in a similar way.

BF

5. ~~A local limit theorem~~ Sufficiency with $0 < \alpha < 1/6$

We now pass to ~~a~~ the local limit theorem. Let the variables X_j possess the bounded continuous density $g(x)$. We introduce some notation. By the letter B we shall denote a bounded function of the parameters considered, not always ~~one and~~ the same. $\delta_0, \delta_1, \cdots; \varepsilon_0, \varepsilon_1, \cdots$ will be small positive constants; $C_0, C_1, C_2, \cdots; c_0, c_1, \cdots$ positive constants. ~~Denote~~ Write

$$(5.1) \qquad \varphi(t) = E\, e^{itX_j} = \int_{-\infty}^{\infty} e^{itu}\, g(u)\, du.$$

The function $|\varphi(t)|^2$ is a nonnegative Fourier transform and therefore (compare [1], p. 20) we have $|\varphi(t)|^2 \in L_1(-\infty, \infty)$, so that

$$(5.2) \qquad \int_{-\infty}^{\infty} |\varphi(t)|^2\, dt < \infty.$$

Hence we have

$$p_{Z_n}(x) = \frac{\sqrt{n}}{2\pi} \int_{-\infty}^{\infty} [\varphi(t)]^n e^{-\sqrt{n}\, itx}\, dt. \tag{5.3}$$

Let (4.1) be ~~fulfilled~~ satisfied. Suppose first that $\alpha < 1/6$. We must prove that $[0, n^{\alpha}/\rho(n)]$ and $[-n^{\alpha}/\rho(n), 0]$ are z.u.l.n.c. We shall study only the first zone, the second one being treated analogously. Take x such that

$$0 \leqq x \leqq \frac{n^{\alpha}}{\rho(n)}\ . \tag{5.4}$$

In view of (4.1) the function $\varphi(t)$ is infinitely differentiable on the whole axis and so, for any $T > 0$, $|t| \leqq T$, and $p > 0$ an integer,

$$\varphi(t) = \varphi(0) + \frac{t\varphi'(0)}{1!} + \cdots + \frac{t^{p-1}}{(p-1)!}\varphi^{(p-1)}(0) + \frac{t^p}{p!}R_p(t), \tag{5.5}$$

where

~~In MS "$R_p(t)$" not defined~~ (MS wrong)

$$|R_p(t)| \leqq 2 \sup_{|t| \leqq T} |\phi^{(p)}(t)|. \tag{5.6}$$

Moreover, $\varphi'(0) = 0$; $\varphi''(0) = -1$. From this it follows that for $|t| \leqq \varepsilon_0$,

$\phi''(0)$

$$(5.7)\qquad \varphi(t) = 1 - \frac{t^2}{2} + Bt^3; \qquad |\varphi(t)| \leqq 1 - \frac{t^2}{4}.$$

Further, in view of the existence of the bounded continuous density $g(x)$, we have, for $|t| > \varepsilon_0$,

$$(5.8)\qquad |\varphi(t)| < 1; \quad \varphi(t) \to 0 \quad \text{as} \quad t \to \pm\infty.$$

Combining (5.2), (5.3), and (5.8) we get

$$(5.9)\qquad p_{Z_n}(x) = \frac{\sqrt{n}}{2\pi}\int_{-\varepsilon_0}^{\varepsilon_0} [\varphi(t)]^n e^{-\sqrt{n}\, itx}\, dt + B e^{-c_0 n}.$$

Put

$$(5.10)\qquad \mu = \frac{1}{2} - \alpha.$$

In view of (5.7), for $n^{-\mu} \leqq |t| \leqq \varepsilon_0$, we ~~get~~ obtain

$$(5.11)\qquad |\varphi(t)|^n \leqq \left(1 - \frac{n^{-2\mu}}{4}\right)^n = B \exp(-c_1 n^{1-2\mu})$$
$$= B \exp(-c_1 n^{2\alpha}).$$

Hence from (5.9) we ~~get~~ obtain

$$(5.12)\qquad p_{Z_n}(x) = \frac{\sqrt{n}}{2\pi}\int_{-n^{-\mu}}^{n^{-\mu}} [\varphi(t)]^n e^{-\sqrt{n}\, itx}\, dt$$
$$+ B \exp(-\delta_0 n^{2\alpha}).$$

6. ~~Continuation of proof~~ Approximation to $\phi(t)$

The function (5.5) is not analytic in general and so the Taylor series for it diverges. We must choose an appropriate approximation to it in the segment $|t| \leqq n^{-\mu}$, that is, choose a convenient p in the formula (5.5). We need ~~the~~ estimates of $\varphi^{(q)}(0)$ for $q \leqq p$. We get

(6.1) $$\varphi^{(q)}(t) = \int_{-\infty}^{\infty} e^{itx} (ix)^q g(x)\, dx.$$

Hence

(6.2) $$|\varphi^{(q)}(t)| \leqq \int_{-\infty}^{\infty} |x|^q g(x)\, dx.$$

Putting $k = (1+2\alpha)/4\alpha$, we obtain from (4.1)

(6.3) $$\int_{-\infty}^{\infty} \exp(|x|^{1/k})\, g(x)dx < \infty,$$

Whence we easily obtain

(6.4) $$|\varphi^{(q)}(t)| = B^q \Gamma(kq).$$

Let us ~~denote~~ write $K(t) = \log \varphi(t)$, where $K(0) = 0$, $|t| \leqq n^{-\mu}$. From (5.12) we ~~get~~ obtain

(6.5) $$p_{Z_n}(x) = \frac{\sqrt{n}}{2\pi} \int_{-n^{-\mu}}^{n^{-\mu}} \exp[nK(t) - \sqrt{n}\, itx]\, dt +$$
$$+ B \exp(-\delta_0 n^{2\alpha}).$$

Moreover, from (5.7) we conclude that, for $|t| \leqq \varepsilon_0$,

(6.6) $$K(t) = B.$$

We need estimates for $K^{(q)}(t_0)$, for $|t_0| \leqq \varepsilon_0/2$.

Putting

$$(6.7) \qquad \tilde{\varphi}(t+t_0) = \varphi(t_0) + \frac{t\varphi'(t_0)}{1!} + \cdots + \frac{t^q\varphi^{(q)}(t_0)}{q!}$$

we ~~get~~ have:

$$(6.8) \qquad K^{(q)}(t_0) = [\log \tilde{\varphi}(t+t_0)]^{(q)}.$$

In fact, to compute $K^{(q)}(t_0) = [\log \varphi(t)]^{(q)}$, with $t=t_0$, we need only the derivatives $\varphi^{(p)}(t_0)$, with $p \leqq q$, and these coincide with $\tilde{\varphi}^{(p)}(t_0)$. If C_ρ is ~~the~~ a circle $|t| \leqq \rho$ where $\tilde{\varphi}(t+t_0)$ has no zeros, then $\log \tilde{\varphi}(t+t_0)$ is analytic for $|t| \leqq \rho$ and

Sign

$$(6.9) \qquad K^{(q)}(t_0) = \frac{q!}{2\pi i} \oint_{C_\rho} \frac{\log \tilde{\varphi}(t+t_0)}{t^{q+1}}\, dt.$$

Consider now (6.7). From (6.4) we have

$$(6.10) \qquad \frac{\varphi^{(p)}(t_0)}{p!} \leqq B \exp[Bp + (k-1)p \log p].$$

We choose now for C_ρ the circle

$(\varphi^{(p)}(t_0))$

$$(6.11) \qquad |t| \leqq \exp[-C_0 - (k-1) \log q].$$

For sufficiently large C_0, we ~~get~~ have inside C_ρ

$$(6.12) \qquad \left| \sum_{p=1}^{q} \frac{\varphi^{(p)}(t_0)}{p!} t^p \right| < \frac{1}{4}$$

as is easily seen from (6.10) and (6.11). Hence, on C_ρ we have

$$(6.13) \qquad \frac{1}{2} < |\tilde{\varphi}(t+t_0)| < \frac{3}{2}$$

solidus

in view of (5.7). Hence, from (6.9), for $|t_0| \leqq \varepsilon_0/2$,

$$(6.14) \qquad K^{(q)}(t_0) = B \exp(Bq + kq \overset{\log}{\cancel{\ln}}\, q).$$

We write now the ~~development~~ expansion (5.5) for the function $K(t)$ instead of $\varphi(t)$ and we put

$$(6.15) \qquad p = m = \left[\frac{n^{2\alpha}}{\rho_1(n)}\right],$$

~~when~~ where the function $\rho_1(n)$ will be fixed later on depending upon $\rho(n)$. Hence, for $|t| \leqq n^{-\mu}$, ~~(compare~~ as in (5.6),

$$(6.16) \qquad \left|\frac{t^m R_m(t)}{m!}\right| = B \exp m[B + (k-1) \overset{\log}{\cancel{\ln}}\, m - \mu \overset{\log}{\cancel{\ln}}\, n] =$$

$$= B \exp m[B - (k-1) \overset{\log}{\cancel{\ln}}\, \rho_1(n)]$$

by definition of m, μ, and $\overset{k}{\cancel{\rho}}$. Moreover, $k = (2\alpha+1)/4\alpha > 1$ because ~~as~~ $\alpha < 1/2$, in fact ~~(even)~~ $\alpha < 1/6$. Hence, for $|t| \leqq n^{-\mu}$,

$$(6.17) \qquad nK(t) = -n\frac{t^2}{2} + n\sum_{r=3}^{m} \Psi_r \frac{t^r}{r!} + B\exp\left(-\delta_1 \frac{n^{2\alpha}}{\rho_1(n)}\right).$$

Re is roman

Now, Re nK(t) $\leqq 0$ for $|t| \leqq n^{-\mu}$. Putting

$$(6.18) \qquad K_3(t) = \sum_{r=3}^{m} \Psi_r \frac{t^r}{r!}$$

we ~~get~~ obtain

$$(6.19) \qquad p_{Z_n}(x) = \frac{\sqrt{n}}{2\pi} \int_{-n^{-\mu}}^{n^{-\mu}} \exp\left[-\frac{nt^2}{2} + nK_3(t) - \sqrt{n}\, itx\right] dt + B \exp\left[-\delta_2 \frac{n^{2\alpha}}{\rho_1(n)}\right].$$

We consider now the entire function

$$(6.20) \qquad \exp\left[nK_3(t)\right] = 1 + \sum_{r=3}^{\infty} \frac{\chi_r t^r}{r!}.$$

With $\alpha < 1/6$, $\mu = 1/2 - \alpha > 1/3$, ~~then~~ for $|t| \leqq n^{-\mu}$, we have

$n\Psi_3\,(t^3/3!) = Bn^{-\varepsilon}$. Taking this into account, and using the estimate, as in ~~(compare~~ (6.14)~~)~~,

$$(6.21) \qquad \Psi_r = K^{(r)}(0) = B \exp\,(Br + kr \log r), \qquad r \geqq 3,$$

we ~~get~~ obtain, after an easy computation,

$$(6.22) \qquad nK_3(t) = B, \quad |t| \leqq 2n^{-\mu},$$

$\oint_{|t|=2n^{-\mu}}$

Sign

$$(6.23)\qquad \chi_r = \frac{r!}{2\pi i}\oint_{|t|=2n^{-\mu}} \exp[nK_3(t)]\frac{dt}{t^{r+1}} = Br!\,2^{-r}n^{r\mu},$$

$$(6.24)\qquad \exp[nK_3(t)] = \sum_{r=3}^{m}\frac{\chi_r t^r}{r!} + B\,2^{-m}, \qquad |t| \le n^{-\mu}.$$

BF

7. Conclusion of the proof

Putting (6.24) into (6.19) and taking (6.15) into account ~~(6.15)~~, we find that

$$(7.1)\qquad p_{Z_n}(x) = \frac{\sqrt{n}}{2\pi}\int_{-n^{-\mu}}^{n^{-\mu}} \exp\left(-\frac{nt^2}{2}\right)\left(1 + \sum_{r=3}^{m}\frac{\chi_r}{r!}t^r\right)$$

$$\cdot \exp(-itx\sqrt{n})\,dt + B\exp\left(-\delta_3\frac{n^{2\alpha}}{\rho_1(n)}\right).$$

To printer: Put on line above if possible. Then delete dot

The next step is to extend the limits of integration in (7.1) from $-\infty$ to ∞. Putting $\xi = t\sqrt{n}$, we obtain

$$(7.2)\qquad p_{Z_n}(x) = \frac{1}{2\pi}\int_{-n^{1/2-\mu}}^{n^{1/2-\mu}} \exp\left(-\frac{\xi^2}{2}\right)$$

$$\cdot\left[1 + \sum_{r=3}^{m}\frac{\chi_r}{r!}\left(\frac{\xi}{\sqrt{n}}\right)^r\right]\exp(-i\xi x)\,d\xi$$

$$+ B\exp\left[-\delta_3\frac{n^{2\alpha}}{\rho_1(n)}\right].$$

Printer: This is 1/2 "one-half minus mu"

2π no root sign

lc en

In order to obtain the required extension of the limit of integration, we estimate, for $3 \leqq r \leqq m$,

$$(7.3) \quad \int_{n^{1/2-\mu}}^{\infty} e^{-\xi^2/2} \frac{\chi_r}{r!} n^{-r/2} \xi^r d\xi = B^r \exp\left(-\frac{1}{4} n^{1-2\mu}\right) \Gamma\left(\frac{r}{2}\right) n^{(\mu-\frac{1}{2})r} .$$

Now

solidus

$$(7.4) \quad \boxed{} \; B^r \Gamma\left(\tfrac{r}{2}\right) n^{(\mu-\frac{1}{2})r} \leqq B \exp\left[Br + \frac{r}{2} \log r - r\left(\frac{1}{2} - \mu\right) \log n\right]$$

$$\leqq B \exp r\left[B + \frac{1}{2} \log m - \left(\frac{1}{2} - \mu\right) \log n\right] = B \exp\left[-\frac{r}{2} \log \rho_1(n)\right].$$

Summing over $r \leqq m$, we obtain from (7.3),

$$(7.5) \quad \sum_{3 \leqq r \leqq m} \int_{n^{1/2-\mu}}^{\infty} e^{-\xi^2/2} \frac{\chi_r}{r!} \left(\frac{1}{\sqrt{n}}\right)^r \xi^r d\xi = B \exp\left(-\delta_4 n^{2\alpha}\right).$$

Of course, the integrals $\int_{-\infty}^{-n^{1/2-\mu}}$ can be subjected to the same treatment, and we get

$$(7.6) \quad \boxed{} \; P_{Z_n}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left(-\frac{\xi^2}{2}\right)\left[1 + \sum_{r=3}^{m} \frac{\chi_r}{r!} \left(\frac{\xi}{\sqrt{n}}\right)^r\right] \exp(-i\xi x)\, d\xi + B \exp\left(-\delta_3 \frac{n^{2\alpha}}{\rho_1(n)}\right).$$

To printer: For command, shift to right and down

We are thus led to the expressions

$$(7.7) \quad n^{-r/2} \int_{-\infty}^{\infty} e^{-\xi^2/2} e^{-i\xi x} \xi^r d\xi = n^{-r/2} H_r^{(0)}(x) e^{-x^2/2}$$

for $3 \leqq r \leqq m$. The polynomials $H_r^{(0)}(x)$ are Hermite polynomials in ax for ~~the~~ suitable $a > 0$. Let now

$$(7.8) \quad |x| \leqq \frac{n^{\alpha}}{\rho(n)} .$$

For $3 \leqq r \leqq C_1$, when C_1 is any constant, we have for (7.7) the value

$$(7.9) \quad B\, e^{-x^2/2} \left[\frac{1}{\rho(n)}\right]^r .$$

$n \leqq C_1$

Summing these for $r = 3,4,\dots; C_1$ we get the estimate

(7.10) $$B\, e^{-x^2/2} \frac{1}{[\rho(n)]^3} .$$

Consider now $C_1 < r \leqq m$. We use the following expression for Hermite's polynomials $H_q(x)$ (see [7], p. 193):

(7.11) $$H_q(x) = q! \sum_{s=0}^{[q/2]} \frac{(-1)^s (2x)^{q-2s}}{s!\,(q-2s)!} .$$

Hence ($x \geqq 0$; 0^0 assumed to be equal to 1),

(7.12) $$|H_q^{(0)}(x)| = B^q q! \max_{s \leqq q/2} \frac{x^{q-2s}}{s!\,(q-2s)!} .$$

Let $s = q\rho$ with $0 \leqq \rho \leqq 1/2$, where $0 \leqq x \leqq n^\alpha/\rho(n)$ and $C_1 < q \leqq m$. Then (7.12) has the value

(7.13) $$B \max_\rho \exp q\{B + (1-2\rho)[\alpha \overset{\log}{\cancel{\ln}}\, n - \overset{\log}{\cancel{\ln}}\, \rho(n)] - \rho \overset{\log}{\cancel{\ln}}\, q - \rho \overset{\log}{\cancel{\ln}}\, \rho - (1-2\rho) \overset{\log}{\cancel{\ln}}\, q - (1-2\rho)\log(1-2\rho)\,\cancel{\ln q}\; \cancel{(1-2\rho)} + \overset{\log}{\cancel{\ln}}\, q\} .$$

Multiplying (7.13) by $n^{-q/2} = \exp[-(q \overset{\log}{\cancel{\ln}}\, n)/2]$ and by

(7.14) $$\frac{X_q}{q!} = Bn^{q\mu} = B \exp(q\mu \overset{\log}{\cancel{\ln}}\, n),$$

we get, after an elementary computation,

(7.15) $$B \max_\rho \exp q[B - (1-2\rho) \overset{\log}{\cancel{\ln}}\, \rho(n) + \rho \overset{\log}{\cancel{\ln}}\, q - \rho 2\alpha \overset{\log}{\cancel{\ln}}\, n].$$

As $\overset{\log}{\cancel{\ln}}\, q \leqq \overset{\log}{\cancel{\ln}}\, m = 2\alpha \overset{\log}{\cancel{\ln}}\, n - \overset{\log}{\cancel{\ln}}\, \rho_1(n) + B$, we can replace (7.15) by

(7.16) $$B \max \exp q[B - (1-2\rho) \overset{\log}{\cancel{\ln}}\, \rho(n) - \rho \overset{\log}{\cancel{\ln}}\, \rho(n)].$$

If $\rho > 1/4$, then $\rho \overset{\log}{\cancel{\ln}}\, \rho(n) > \overset{\log}{\cancel{\ln}}\, \rho(n)/4$; if $\rho \leqq 1/4$, then $(1-2\rho) \overset{\log}{\cancel{\ln}}\, \rho(n) \geqq \overset{\log}{\cancel{\ln}}\, \rho(n)/2$. Hence (7.16) can be replaced by lc pee

(7.17) $B[\rho(n)]^{-\varepsilon_2 q}$, ~~for~~ $C_1 < q \leq m$.

Inserting this into (7.6) and summing over q, we ~~get:~~ obtain

(7.18) $$P_{Z_n}(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \left[1 + \frac{B}{\rho(n)}\right] + B \exp\left[-\delta_3 \frac{n^{2\alpha}}{\rho_1(n)}\right].$$

Note that C_1 is supposed to be chosen $> 1/\varepsilon_2$. The relation (7.18) holds for $0 \leqq x \leqq n^\alpha/\rho(n)$. We take now $\rho_1(n) = \rho(n)$, from which

(7.19) $$P_{Z_n}(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \left[1 + \frac{B}{\rho(n)}\right], \quad 0 \leqq x \leqq \frac{n^\alpha}{\rho(n)},$$

lc pee

which was to be proved. The case where $-n^\alpha/\rho(n) \leqq x \leqq 0$ is treated similarly.

8. Proof of sufficiency when $1/6 \leqq \alpha < 1/2$

We pass now to the case $1/6 \leqq \alpha < 1/2$. The principal scheme of reasoning remains the same as in the previous case, for which $\alpha < 1/6$; we shall indicate only the points where it differs substantially from the previous scheme.

Let $1/6 \leqq \alpha < 1/2$ and assume ~~the~~ condition (4.1) holds. In view of (5.7) and (4.1) we know that $K(t) = \log \varphi(t)$ possesses all derivatives for $|t| \leqq \varepsilon_0$, and that $\Psi_r = K^{(r)}(0)$, when multiplied into a suitable power of $i = \sqrt{-1}$, becomes the rth cumulant of X_j. As is well known ([9], pp. 61-63), the rth cumulant depends only upon the moments up to the rth, and conversely. If the third, fourth, ..., ~~3d, 4th,~~ rth cumulant vanish, while $EX_j = 0$ and $D(X_j) = 1$, then all the moments up to the rth are the moments of the normal law $N(0,1)$. Consider the sequence of the critical numbers (2.3) and suppose that

(8.1) $$\frac{1}{2}\,\frac{s+1}{s+3} \leqq \alpha < \frac{1}{2}\,\frac{s+2}{s+4}.$$

Consider the numbers $\Psi_3, \Psi_4, \ldots, \Psi_{s+3}$. If all these numbers vanish (and so the first $s + 3$ moments of X_j are normal), we proceed as in the previous case. We put $\mu = 1/2 - \alpha$, and hence, $\mu \leq 1/3$, and carry out the above computations. Instead of (6.17) we have now

$$(8.2) \qquad nK(t) = -\frac{nt^2}{2} + n \sum_{r=s+4}^{m} \Psi_r \frac{t^r}{r!} + B \exp\left[-\delta_1 \frac{n^{2\alpha}}{\rho(n)}\right]$$

for $|t| \leq n^{-\mu}$. Since $\mu = 1/2 - \alpha \geq$ ~~$\frac{1}{s+4}$~~ $1/(s+3)$, we have

$$(8.3) \qquad n\Psi_{s+3} \frac{t^{s+4}}{(s+4)!} = Bn^{1-\frac{s+4}{s+3}} = Bn^{-\frac{1}{s+3}} .$$

Reasoning now as in section 7, we obtain the local limit theorem for $0 \leq x \leq n^{\alpha}/\rho(n)$ and $-n^{\alpha}/\rho(n) \leq x \leq 0$.

9. Proof of necessity ~~in case~~ when $1/6 \leq \alpha < 1/2$.

Suppose now that not all the numbers $\Psi_3, \Psi_4, \ldots, \Psi_{s+3}$ vanish; let Ψ_{s_0+3} be the first nonvanishing number for $s_0 \leq s$. We then set

$$(9.1) \qquad \mu = \frac{1}{s_0 + 3} .$$

We can show that in this case there will be no local normal convergence even in the zone $[0, n^{1/2-\mu}]$ which is obviously smaller than the zone $[0, n^{\alpha}\rho(n)]$. We choose $m = [n^{1-2\mu}/C_2]$ with a sufficiently large C_2. We have then, proceeding as in sections 6 and 7,

α

le pee

$$(9.2) \qquad P_{Z_n}(x) = \frac{\sqrt{n}}{2\pi} \int_{-n^{-\mu}}^{n^{-\mu}} \exp\left[-\frac{nt^2}{2} + K_{s_0+3}(t) - i\sqrt{n}\,tx\right]dt$$

$$+ B \exp(-\delta_5 n^{1-2\mu}).$$

In this case

(9.3) $$K_{s_0+3}(t) = \sum_{r=s_0+3}^{m} \frac{\Psi_r t^r}{r!} .$$

Further,

(9.4) $$\exp\,[nK_{s_0+3}(t)] = 1 + \sum_{r=s_0+3}^{\infty} \frac{X_r t^r}{r!} .$$

It is important to notice that

(9.5) $$X_{s_0+3} = n\Psi_{s_0+3} , \quad (\Psi_{s_0+3} \neq 0) .$$

Taking now

(9.6) $$x_0 = \eta n^{1/2-\mu} , \quad \eta > 0$$

with a sufficiently small η, and proceeding as in section 7, we obtain, after some computations similar to that of section 7,

$le\ pee^{-x_0^2}$

(9.7) $$P_{Z_n}(x) = \frac{1}{\sqrt{2\pi}} e^{-x_0^2/2} + \frac{1}{2\pi} \frac{X_{s_0+3}}{(s_0+3)!} H_{s_0+3}(x_0) e^{-x_0^2/2} \frac{1}{(\sqrt{n})^{s_0+3}} + B e^{-x_0^2/2} \eta^{s_0+4} ,$$

$e^{-x_0^2/2}$

for a sufficiently small $\eta > 0$. Taking (9.5) into account ~~(9.5)~~ we see that the second term on the right ~~hand~~ side is of the form

$e^{-x_0^2/2}$

(9.8) $$a_0 \Psi_{s_0+3} \frac{1}{(s_0+3)!} e^{-x_0^2/2} \eta^{s_0+3} [1 + o(1)]$$

as $n \to \infty$, where $a_0 \neq 0$. Hence, there is no local normal convergence in the zone $[0, n^{1/2-\mu}]$ and likewise in the zone $[0, n^{\alpha}\rho(n)]$.

Hence, if there is ~~the~~ uniform local normal convergence in the zones $[0, n^{\alpha}]$ and $[-n^{\alpha}, 0]$ for all $\alpha < 1/2$, then all Ψ_r must vanish and so all the X_j's are normal.

10. An integral theorem uniform for the whole x-axis

There is an interesting class of probability densities $g(x)$ for which an integral limit theorem for the normal sum Z_n holds for the whole x-axis.

Consider the class of all even[1] continuous probability densities $g(x)$ such that for $x \geqq 1$

$$(10.1) \qquad P\{X_1 > x\} = \int_x^{\infty} g(u)du = \frac{A_a}{x^a} + \frac{A_{a+1}}{x^{a+1}} + \dots + \frac{A_{4a+5}}{x^{4a+5}} + O\left(\frac{1}{x^{4a+5+\varepsilon}}\right)$$

cap/ nok

where $a \geqq 3$, a being an integer, A_j being constants. Let $X_1, X_2, \dots, X_n$ be ~~the~~ random variables with ~~the~~ probability density $g(x)$ of this class. Then $E(X_j) = 0$. ~~Denote~~ Write $D(X_j) = \sigma^2$, $Z_n = (X_1 + \dots + X_n)/\sigma\sqrt{n}$.

solidus

<u>Theorem 6</u>. <u>For</u> $x \geqq 1$, <u>and as</u> $n \to \infty$ <u>we have uniformly with respect to</u> x,

$$(10.2) \qquad \frac{P\{Z_n > x\}}{\frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-u^2/2}du + r(x, \sqrt{n})} \to 1,$$

<u>where</u> $r(x, \sqrt{n})$ <u>is a rational function of both variables determined by the coefficients</u> $A_a, \dots, A_{4a+5}$ <u>in</u> (10.1). <u>Moreover, for</u> $x \geqq n^{3/2 + 1/a} \log n$ <u>we have</u> lc

[1] Note that the evenness condition is assumed for simplicity only.

(10.3) $$r(x,\sqrt{n}) \sim nP\{X_1 > \sigma x\sqrt{n}\} = n\int_{\sigma x\sqrt{n}}^{\infty} g(u)\,du.$$

Of course, $g(x)$ being even, an analogous relation holds for $x \leqq -1$, while for $-1 < x < 1$ the classical normal convergence theorem acts.

Simple examples of theorem 6 are given by ~~the~~ rational densities. For instance, if $g(x) = 2/[\pi(x^2+1)^2]$, so that $D(X_j) = 1$, we have for $x \geqq 1$, $n \to \infty$,

(10.4) $$\frac{P\{Z_n > x\}}{\frac{1}{\sqrt{2\pi}}\int_x^{\infty} e^{-u^2/2}\,du + \frac{2}{3\pi\sqrt{n}}\,\frac{1}{x^3}} \longrightarrow 1,$$

and

(10.5) $$\frac{2}{3\pi\sqrt{n}}\,\frac{1}{x^3} \sim nP\{X_1 > x\sqrt{n}\}$$

for $x \geqq n^{3/2+1/4}\log n$.

We shall indicate here briefly the principal points of the corresponding proof. We take $\sigma = 1$. The case

(10.6) $$x \geqq n^{3/2+1/a}\log n$$

is treated in an elementary manner. If $y > n$, then the event $S_n > y$ implies at least one of the events $X_j > y/n$. Denote by $H_{k,\alpha_1,\alpha_2,\ldots,\alpha_k}$, ~~for~~ $k \leqq n$, the hypothesis that $X_{\alpha_1} > y/n$, $X_{\alpha_2} > y/n, \ldots, X_{\alpha_k} > y/n$, while this is not true for any other X_j. Hence

(10.7) $$\begin{aligned} P\{S_n > y\} &= \sum_{(\alpha_1)} P\{H_{1,\alpha_1}\}P\{S_n > y \mid H_{1,\alpha_1}\} \\ &+ \sum_{k\geqq 2}\ \sum_{(\alpha_1,\ldots,\alpha_k)} P\{H_{k,\alpha_1,\ldots,\alpha_k}\}P\{S_n > y \mid H_{k,\alpha_1,\ldots,\alpha_k}\}. \end{aligned}$$

Do not break

the summation being extended to all ordered sets of distinct numbers $\alpha_1 < \alpha_2 < \ldots < \alpha_k$. Now, in view of (10.1) it is easy to compute that for

$$y \geq n^{2+1/a} \log n,$$

the double sum ~~at~~ on the right ~~hand~~ side of (10.7) is of order smaller than the first ~~one~~ term, so

$$P\{S_n > y\} \sim \sum_{(\alpha_1)} P\{H_{1,\alpha_1}\} \; P\{S_n > y \mid H_{1,\alpha_1}\}$$

(10.8)

$$= nP\{H_{1,1}\} \; P\{S_n > y \mid H_{1,1}\}.$$

Moreover, it is easy to see that

$$P\{S_n > y \mid H_{1,1}\} \sim P\{X_1 > y \mid X_1 > \frac{y}{n}\}. \tag{10.9}$$

Inserting this into (10.8) we get:

$$P\{S_n > y\} \sim nP\{X_1 > y\} \qquad y > n^{2+1/a} \log n$$

or

$$P\{Z_n > x\} \sim nP\{X_1 > x\sqrt{n}\} \tag{10.10}$$

for

$$x \geq n^{3/2 + 1/a} \log n. \tag{10.11}$$

We must now investigate the behavior of $P\{Z_n > x\}$ for $1 \leq x < n^{3/2 + 1/a} \log n$. It is possible to do so by the method expounded in sections 6 and 7.

Let $\varphi(t)$ be a characteristic function. A function $\gamma(t)$ will be called ~~the~~ a radial continuation of $\varphi(t)$ for the ray $t \geq 0$, if it is defined in some neighborhood of $t = 0$ and coincides with $\varphi(t)$ in this neighborhood for $t \geq 0$ (a radial continuation for $t \leq 0$ is similarly defined). For instance, the function $\varphi(t) = e^{-|t|}(|t| + 1)$, corresponding to the density $g(x) = 2/\pi\,(x^2 + 1)^2$, has two radial continuations: $\gamma(t) = e^{-t}(t + 1)$ for $t \geq 0$, and $\gamma(t) = e^{t}(-t + 1)$ for $t \leq 0$. Both continuations are entire functions (though different ones). Note that they are not even, while $\varphi(t)$ is even.

11. Continuation of the proof

From (10.1) we deduce by integration by parts that $\varphi(t)$ has a radial continuation $\gamma(t)$, coinciding with it for $t \geq 0$, which is differentiable at least $b = 4a + 2$ times. We proceed now to calculate $p_{Z_n}(x)$ (the integral theorem can be obtained later by integration). As $g(x)$ is even, $\varphi(t)$ is real and we obtain

$$p_{Z_n}(x) = \frac{\sqrt{n}}{\pi}\,\mathrm{Re}\int_0^{\infty}[\varphi(t)]^n e^{-\sqrt{n}\,itx}\,dt. \tag{11.1}$$

Applying the reasoning of section 5 [compare (5.9)] we get:

$$\begin{aligned} p_{Z_n}(x) &= \frac{\sqrt{n}}{\pi}\,\mathrm{Re}\int_0^{\varepsilon_0}[\varphi(t)]^n e^{-\sqrt{n}\,itx}\,dt + Be^{-c_0 n} \\ &= \frac{\sqrt{n}}{\pi}\,\mathrm{Re}\int_0^{\varepsilon_0}[\gamma(t)]^n e^{-\sqrt{n}\,itx}\,dt + Be^{-c_0 n}. \end{aligned} \tag{11.2}$$

Further

$$\varphi(t) = \gamma(t) = 1 - \frac{t^2}{2} + O(t^3), \quad (0 \leq t \leq \varepsilon_0). \tag{11.3}$$

From this,

$$p_{Z_n}(x) = \frac{\sqrt{n}}{\pi}\,\mathrm{Re}\int_0^{\ln n/\sqrt{n}}[\gamma(t)]^n e^{-\sqrt{n}\,itx}\,dt + Be^{-\delta_0 \ln^2 n}. \tag{11.4}$$

Also

$$\gamma(t) = \gamma_0(t) + B\,|t|^b; \quad \gamma_0(t) = 1 - \frac{t^2}{2} + \sum_{q=3}^{b-1}\frac{\gamma^{(q)}(0)}{q!}t^q + Bt^b \tag{11.5}$$

for $|t| \leq \varepsilon_0$. For $|t| \leq \overset{\log}{\cancel{\ln}}\, n/\sqrt{n}$, we have

$$\gamma(t) = \gamma_0(t) + B_\varepsilon n^{-b/2+\varepsilon} \tag{11.6}$$

for any $\varepsilon > 0$, B_ε depending upon ε. Further

$$K(t) = \overset{\log}{\cancel{\ln}}\,\gamma(t) = \overset{\log}{\cancel{\ln}}\,\gamma_0(t) + B_\varepsilon n^{-b/2+\varepsilon}, \quad \left(0 \leq t \leq \frac{\overset{\log}{\cancel{\ln}}\, n}{\sqrt{n}}\right). \tag{11.7}$$

Further, for $0 \leq t \leq \overset{\log}{\cancel{\ln}}\, n/\sqrt{n}$,

$$\overset{\log}{\cancel{\ln}}\,\gamma_0(t) = -\frac{t^2}{2} + \sum_{q=3}^{b-1} \Psi_q \frac{t^q}{q!} + Bt^b. \tag{11.8}$$

For $0 \leq t \leq \overset{\log}{\cancel{\ln}}\, n/\sqrt{n}$, we have $Bnt^b = B_\varepsilon n^{-b/2+1+\varepsilon}$.

Thus

$$p_{Z_n}(x) = \frac{\sqrt{n}}{\pi}\,\mathrm{Re}\int_0^{\overset{\log}{\cancel{\ln}}\, n/\sqrt{n}} \exp n\left(-\frac{t^2}{2} + \sum_{q=3}^{b-1} \Psi_q \frac{t^q}{q!}\right. \left. - \sqrt{n}\, itx\right) dt + B_\varepsilon n^{-b/2+1+\varepsilon}. \tag{11.9}$$

If we write ~~denote~~ $K_3(t) = \sum_{q=3}^{b-1} \Psi_q\, t^q/q!$ we obtain

$$nK_3(t) = B_\varepsilon n^{-1/2+\varepsilon}, \quad 0 \leq t \leq \frac{\overset{\log}{\cancel{\ln}}\, n}{\sqrt{n}}.$$

Hence

$$\exp\,[nK_3(t)] = 1 + K_4(t,n) + B_\varepsilon n^{-b/2+1+\varepsilon}, \tag{11.10}$$

where

$$K_4(t,n) = \sum_{q=1}^{b} \frac{[nK_3(t)]^q}{q!}. \tag{11.11}$$

Hence, taking $t = \xi/\sqrt{n}$, we get

$$p_{Z_n}(x) = \frac{1}{\pi}\,\mathrm{Re}\int_0^{\log n} \exp\left(-\frac{\xi^2}{2}\right)\left[1 + K_4\left(\frac{\xi}{\sqrt{n}}, n\right)\right]\exp(-i\xi x)\,d\xi + B_\varepsilon n^{-b/2+1+\varepsilon}. \tag{11.12}$$

space

Extending the integration limit to $\xi = \infty$ (the error is estimated trivially),

$$\boxed{}\; p_{Z_n}(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} + \frac{1}{\pi} \text{ non } \mathrm{Re}\int_0^\infty \exp\left(-\frac{\xi^2}{2}\right) K_4\left(\frac{\xi}{\sqrt{n}}, n\right) \exp(-i\xi x)\,d\xi + B_\varepsilon n^{-b/2+1+\varepsilon}. \tag{11.13}$$

The function $K_4(\xi/\sqrt{n}, n)$ is a polynomial with respect to $\xi/\sqrt{n}$. The evaluation of (11.13) is thus reduced to the evaluation of the integrals

$$E(x,r) = \text{non } \mathrm{Re}\int_0^\infty e^{-\xi^2/2} e^{i\xi x} \xi^r d\xi \tag{11.14}$$

for large values of x. If r is an even number, $E(x,r) = (1/2)\int_{-\infty}^{\infty} e^{-\xi^2/2} e^{i\xi x} \xi^r d\xi$ (11.15) (display) is easily expressed ~~through~~ in terms of $e^{-x^2/2}$ and Hermite's polynomials [compare (7.7)]. If r is odd, there is apparently no expression ~~through~~ in terms of the elementary functions, but, for large values of x, (11.14) can be easily evaluated by integration by parts. Thus, for r = 1, we have $E(x,r) \sim 1/x^2$.

We must now show that the formula (11.13) holds for the values of x satisfying

$$1 \leq x \leq n^{3/2 + 1/a} \log n. \tag{11.16}$$

Comparing $nP\{X_1 > x\sqrt{n}\}$ [compare(10.11)] to the remainder term $B_\varepsilon n^{-b/2+1+\varepsilon}$ in (11.13) we deduce that $n^{-b/2+1+\varepsilon}$ must be smaller than $n(n^{2+1/a+\varepsilon})^{-a} = n^{-2a-\varepsilon'}$; from this, $b/2 - 1 > 2a$ or $b > 4a + 2$. Under this condition we obtain the local theorem valid up to $n^{3/2+1/a+\varepsilon}$ uniformly and hence, by integration, the integral theorem. The relation (10.11) enables us to obtain it for the whole x-axis.

12. Concluding remarks

The condition of evenness was assumed only to simplify the final formulas; if it is not fulfilled, (11.13) will only involve $\int_{-\infty}^{0}$ in ~~addition~~ to $\int_{0}^{\infty}$. Moreover, the analogous limit theorem on the whole axis can be obtained for the variables X_j such that

$$(12.1) \qquad P\{X_j > x\} = \int_{\nu=a}^{a_1} \frac{dG(\nu)}{x^\nu} + O\left(\frac{1}{x^{a_1+\varepsilon}}\right), \quad \{x \geq 1\}$$

where $a_1 > 4a + 5$ and $G(\nu)$ is a function of bounded variation. ~~The~~ A similar relation must hold for negative values of x.

The new approach expounded here is applicable also to independent variables which are not identically distributed and to the investigation of nonnormal convergence.

The asymptotic behavior of the ~~big~~ large deviations of order statistics can be also studied by this method.

REFERENCES

[1] S. Bochner and K. Chandrasekharan, Fourier Transforms, Princeton, Princeton University Press, 1949.

[~~2~~ 8] B.V. Gnedenko and A.N. Kolmogorov, Limiting Distributions for Sums of Independent Random Variables, Cambridge, Mass., Addison-Wesley, 1954.

[~~3~~ 2] H. Chernoff, "Large sample theory: parametric case," Ann. Math. Statist., Vol. 27 (1956), pp. 1-22.

[~~4~~ 3] H. Cramér, "Sur un nouveau théoreme-limite de la théorie des probabilitiés," Actualités ~~Actuel.~~ Sci. Ind., No. 736, ~~Paris~~ (1938).

[~~5~~ 4] H. Daniels, "Saddlepoint approximations in statistics," Ann. Math. Statist., Vol. 25 (1954), pp. 631-650.

[~~6~~ 5] R.L. Dobrushin, "A general formulation of the fundamental theorem of Shannon," Uspekhi Mat. Nauk, Vol. 6 (1959), ~~(90)~~, pp. 3-104. ~~(1959)~~ ~~(In Russian.)~~

[~~7~~ 6] A. Erdélyi, Higher Transcental Functions, Vol. ~~II~~ 2, New York, McGraw-Hill, 1953.

[~~8~~ 7] W. Feller, "Generalization of a probability theorem of Cramér," Trans. Amer. Math. Soc., Vol. 54 (1943), pp. 361-372.

[9] M. Kendall, The Advanced Theory of Statistics, Vol. 1, London, Griffin, 1952.

[10] A. I. Khintchine, "Über einer neuen Grenzwertsatz der Wahrscheinlichkeitsrechnung," Math. Ann., Vol. 101 (1929), pp. 745-752.

[11] Yu. V. Linnik, Markov Chains in the Analytic Arithmetic of Quaternions and Matrices, Vestnik Leningrad ~~Government~~ Univ~~ersity,~~ No. 13 (1956), pp. 63-68. (In Russian.)

[12] U.V. Petrov, "A generalization of a limit theorem of Cramér," Uspekhi Mat. Na~~y~~uk, Vol. 9 ~~No. 4~~ (1954), pp. 196-202. (In Russian.)

[13] V. Richter, "A local limit theorem for large deviations," Dokl~~ady~~ Akad. Nauk SSSR, ~~U.S.S.R.,~~ Vol. 115 ~~No. 1~~ (1957), pp. 53-56. (In Russian.)

[14] V. Richter, "Local limit theorems for large deviations," Teor. Veroyatnost. i Primenen., ~~Theory of Probabilities and its Applications,~~ Vol. 2 (1957), ~~No. 2,~~ pp. 214-229. (In Russian.)

[15] V. Richter, "Multidimensional local limit theorems for large deviations," ibid., ~~Theory of Probabilities and its Applications,~~ Vol. 3 (1958), ~~No. 1,~~ pp. 107-114. (In Russian.)

[16] N.V. Smirnov, "On the probabilities of large deviations," Mat. Sb., ~~Math. Sbornik,~~ Vol 40 (1933), pp. 443-454. (In Russian.)

[17] L. Treloar, Physics of Rubber Elasticity, Oxford, Oxford University Press, 1949.

[1~~8~~9] J. Wolfowitz, "Information ~~for~~ theory for mathematicians," Ann. Math. Statist., Vol. 29 (1958), pp. 351-356.

[1~~9~~8] M.V. Volkenshtain and O.B. Ptitsin, "Statistical physics of linear polymer chains," Uspekhi Fiz. ~~Phys.~~ Nauk, ~~Hayk,~~ Vol. 49 ~~No. 2~~ (1953), pp. 501-568. (In Russian.)

Continuity and Conditions of Gelder for Selected Functions of Stationary Gauss Processes. Yu. K. Belayev

НЕПРЕРЫВНОСТЬ И УСЛОВИЯ ГЕЛЬДЕРА ДЛЯ ВЫБОРОЧНЫХ ФУНКЦИЙ СТАЦИОНАРНЫХ ГАУССОВСКИХ ПРОЦЕССОВ

Ю.К.БЕЛЯЕВ

В настоящей работе рассмотрен ряд вопросов, связанных с локальными свойствами выборочных функций стационарных стохастически непрерывных сепарабельных гауссовских процессов. Слова стохастически непрерывный и сепарабельный будут опускаться. Изложенные ниже результаты были опубликованы ранее, см.[I], без доказательств, В § I показано, что для стационарных гауссовских процессов имеет место альтернатива: или почти все выборочные функции непрерывны, или почти все выборочные функции неограничены на любом интервале конечной длины. Гипотеза о существовании такой альтернативы давно высказывалась А.Н.Колмогоровым, см. также [2]. В § 2 достаточные условия для непрерывности выборочных функций, полученные Хантом [3], сформулированы в терминах корреляционных функций,формулы (I2) и (I3). В этом же параграфе построены примеры гауссовских процессов с вероятностью единица неограниченных на интервалах конечной длины. Эти примеры показывают, что условия (I2) и (I3) нельзя усилить существенным образом. Достаточные условия для того, чтобы почти все выборочные функции удовлетворяли условию Гельдера рассматриваются в § 3. Автор выражает глубокую благодарность А.Н.Колмогорову за постановку задач.

§ I Альтернатива

Пусть $\xi(t)$ стационарный гауссовский процесс с непрерывной корреляционной функцией, принимающий вещественные значения.

Основным результатом этого параграфа является следующая

ТЕОРЕМА I. Для любого гауссовского стохастическинепрерывного стационарного процесса $\xi(t)$ справедлива следующая альтернатива: либо с вероятностью единица выборочные функции $\xi(t)$ непрерывны, либо с вероятностью единица они неограничены на любом конечном интервале.

Для доказательства теоремы нам понадобятся две леммы.

ЛЕММА I. Пусть $\zeta(t)$ случайный процесс такой ,что на интервале $\Delta = (t', t'')$

$$P\{\sup_{t\in\Delta} \zeta(t) > a\} > p.$$

Тогда существует конечное множество точек $S_n = \{t_1, \dots, t_n\}$, что

$$P\{\max_{t_i \in S_n} \zeta(t_i) > a\} > p.$$

ЛЕММА 2. Пусть $\zeta_1(t), \dots, \zeta_m(t), \dots, M\zeta_m(t)=0$, последовательность взаимно независимых стационарных гауссовских процессов, у которых для некоторого интервала $\Delta = (t'_1, t'_2)$ и чисел $\varepsilon > 0$, $\delta > 0$

$$P\{\sup_{t\in\Delta} \zeta_k(t) > a\} > p, \quad \sup_{t\in\Delta} P\{\sum_{i=1}^{\infty} \zeta_i(t) < -\delta\} < \frac{\varepsilon}{2m},$$

где m находится из условия $(1-p)^m < \frac{\varepsilon}{2}$.
Тогда

$$P\{\sup_{t\in\Delta} \sum_{i=1}^{m} \zeta_i(t) > a - \delta\} > 1 - \varepsilon.$$

Доказательство этих вспомогательных лемм элементарное и мы его опускаем.

Как известно, модулем непрерывности функции $f(t)$ на отрезке Δ называется функция

$$\omega_f(\Delta,\delta) = \sup_{t',t''\in\Delta,\, |t'-t''|<\delta} |f(t') - f(t'')|.$$

В том случае, когда на интервале Δ почти все выборочные функции непрерывны,

$$\lim_{\delta\downarrow 0} P\{\omega_\xi(\Delta,\delta) > \varepsilon\} = 0$$

для любого $\varepsilon > 0$. В таких обозначениях нам достаточно показать, что если

$$\lim_{\delta\downarrow 0} P\{\omega_\xi(\Delta',\delta) > b\} > p' \qquad \text{(I)}$$

для некоторых чисел $b = 2a > 0$, $p' > 0$ и интервала Δ', то для любого интервала Δ и любого $N > 0$

$$P\{\sup_{t\in\Delta} |\xi(t)| > N\} = 1.$$

Если выполнено неравенство (I), то используя стационарность процесса $\xi(t)$ можно показать, что для любого интервала $\Delta = (t_1, t_1')$ при некотором $p = p(t_1' - t_1) > 0$

$$\lim_{\delta\downarrow 0} P\{\omega_\xi(\Delta,\delta) > b\} > p. \qquad \text{(I}'\text{)}$$

Заметим теперь, что если для некоторого интервала $\Delta = (t_1, t_1')$, $b = 2a > 0$, $p > 0$, $P\{\omega_\xi(\Delta,\delta) > 2a\} > p$, то

$$P\{\sup_{t\in\Delta} |\xi(t) - \xi(t_1)| > a\} > p.$$

Учитывая симметрию гауссовских распределений отсюда получаем, что

$$P\{\sup_{t\in\Delta}[\xi(t)-\xi(t_1)]>a\}>\frac{p}{2}.$$

Из (I′) вытекает, что спектр процесса $\xi(t)$ неограничен, так как в противном случае почти все выборочные функции были бы целыми аналитическими [4]. Поэтому (I′) имеет место при любом $L>0$ для процесса

$$\xi_L(t)=\int\limits_{|\lambda|\geq L}e^{it\lambda}\phi(d\lambda).$$

Здесь $\phi(d\lambda)$ спектральная случайная мера, соответствующая процессу $\xi(t)$. Таким образом, для процесса $\xi_L(t)$

$$P\{\sup_{t\in\Delta}[\xi_L(t)-\xi_L(t_1)]>a\}>\frac{p}{2}. \qquad (2)$$

Пусть нам заданы $\varepsilon>0$, $\delta>0$. Выберем сначала отрезок $\Delta=(t_1,t')$ столь малым, а затем L столь большим, что

$$\inf_{t\in\Delta}P\{\xi(t)-\xi(t_1)>-\delta\}>1-\varepsilon,\ \inf_{t\in\Delta}P\{\xi_L(t)-\xi_L(t_1)>-\delta\}>1-\frac{\varepsilon}{2m}, \qquad (3)$$

где m такое целое число, что

$$\left(1-\tfrac{1}{2}p\right)^m<\frac{\varepsilon}{2}. \qquad (4)$$

Из леммы I и (2) получаем, что для некоторого конечного множества $T_1\subset\Delta$

$$P\{\max_{t_i\in T_1}[\xi_L(t_i)-\xi_L(t_1)]>a\}>\frac{p}{2}.$$

Но

$$\xi_L(t)=\underset{N_1\to\infty}{\mathrm{l.i.m.}}{}^{x)}\int\limits_{N_1>|\lambda|\geq L}e^{it\lambda}\phi(d\lambda)$$

---

x/ l.i.m. означает сходимость в среднем квадратичном.

поэтому найдется такое достаточно большое N_1, что для

$$\xi_1(t) = \int_{N_1 > |\lambda| \geq L} e^{it\lambda} \phi(d\lambda)$$

$$P\{ \max_{t_i \in T_1} [\xi_1(t_i) - \xi_1(t_1)] > a \} > \frac{p}{2}.$$

Так как у процесса

$$\zeta_1(t) = \int_{|\lambda| \geq N_1} e^{it\lambda} \phi(d\lambda)$$

$$\lim_{\delta \downarrow 0} P\{ \omega_{\zeta_1}(\Delta, \delta) > 2a \} > p,$$

то таким же способом находим $N_2 > N_1$ и конечное множество $T_2 \subset \Delta$, что для

$$\xi_2(t) = \int_{N_2 > |\lambda| \geq N_1} e^{it\lambda} \phi(d\lambda)$$

$$P\{ \max_{t_i \in T_2} [\xi_2(t_i) - \xi_2(t_1)] > a \} > \frac{p}{2}.$$

Поступая аналогичным образом и далее, мы строим последовательность взаимно независимых стационарных гауссовских процессов

$$\xi_k(t) = \int_{N_k > |\lambda| \geq N_{k-1}} e^{it\lambda} \phi(d\lambda), \quad k = 3, 4, \ldots,$$

для которых

$$P\{ \sup_{t \in \Delta} [\xi_k(t) - \xi_k(t_1)] > a \} > \frac{p}{2}.$$

Если обозначить $\zeta_k(t) = \xi_k(t) - \xi_k(t_1), k=1,\ldots$, то учитывая (3), (4), и применяя лемму 2, легко получаем, что

$$P\{ \sup_{t \in \Delta} [\xi(t) - \xi(t_1)] > a - 2\delta \} > 1 - 2\varepsilon.$$

Так как $\varepsilon > 0$, $\delta > 0$ и отрезок Δ можно выбирать сколь угодно малыми, то отсюда в свою очередь следует, что для любого отрезка $\Delta = (t_1, t_1')$

$$P\{ \sup_{t\in\Delta} [\xi(t) - \xi(t_1)] > a \} = 1. \qquad (5)$$

На этом заканчивается первый этап доказательства теоремы.

Покажем теперь, что из (5) следует

$$P\{ \sup_{t\in\Delta} [\xi(t) - \xi(t_1)] > na \} = 1, \quad n = 2, 3, \dots \qquad (6)$$

Проведем рассуждения для $n = 2$. Итак, пусть (5) выполнено. Снова зададим сколь угодно малые числа $\varepsilon > 0$, $\delta > 0$. Из (5) и леммы I вытекает, что существует такое конечное множество точек $S \subset \Delta$, для которого

$$P\{ \max_{t_i \in S} [\xi(t_i) - \xi(t_1)] > a \} > 1 - \frac{\varepsilon}{4}.$$

Так как

$$\xi(t) = \underset{M\to\infty}{\mathrm{l.i.m.}} \int_{|\lambda|<M} e^{it\lambda} \Phi(d\lambda),$$

то можно выбирать $M > 0$ столь большим, что для

$$\eta_1(t) = \int_{|\lambda|<M} e^{it\lambda} \Phi(d\lambda)$$

$$P\{ \max_{t_i \in S} [\eta_1(t_i) - \eta_1(t_1)] > a \} > 1 - \frac{\varepsilon}{4}, \qquad (7)$$

а для

$$\eta_2(t) = \int_{|\lambda|\geq M} e^{it\lambda} \Phi(d\lambda)$$

$$\inf_{t\in\Delta} P\{ \eta_2(t) - \eta_2(t_1) > -\delta \} > 1 - \frac{\varepsilon}{2}. \qquad (8)$$

Процесс $\eta_1(t)$ имеет ограниченный спектр. Следовательно, почти все выборочные функции этого процесса непрерывные. Поэтому около каждой точки $t_k \in S' = \{t_2, \dots, t_N\}$, $t_N < t_1'$, можно выделить столь малый отрезок $\Delta_k = (t_k - \delta', t_k + \delta') \subset \Delta$, что в том случае, когда $\eta_1(t_k) - \eta_1(t_1) > a$, то и всюду на отрезке Δ_k $\eta_1(t) - \eta_1(t_1) > a$ для всех k одновременно с вероятностью большей $1 - \frac{\varepsilon}{4}$. Учитывая (7) отсюда получаем, что

$$P\Big\{\bigcup_{k=2}^{N}\Big[\inf_{t\in\Delta_k}(\eta_1(t) - \eta_1(t_1)) > a\Big]\Big\} > 1 - \frac{\varepsilon}{2}. \qquad (9)$$

Рассмотрим теперь процесс $\eta_2(t)$. Повторив рассуждения первой части теоремы, можно показать, что для отрезков Δ_k $k=2,\dots,N$

$$P\{\sup_{t\in\Delta_k}[\eta_2(t) - \eta_2(t_k - \delta')] > a\} = 1.$$

Из этого равенства и из (8), для события A_k, состоящего в том, что $\sup_{t\in\Delta_k}[\eta_2(t) - \eta_2(t_1)] > a - \delta$,

имеем

$$P\{A_k\} \geq P\{\sup_{t\in\Delta_k}[\eta_2(t) - \eta_2(t_k - \delta')] > a\} -$$
$$- P\{\eta_2(t_k - \delta') - \eta_2(t_1) < -\delta\} \geq 1 - \frac{\varepsilon}{2}.$$

Пусть B_k событие, состоящее в том, что для $2 \leq i < k$

$$\inf_{t\in\Delta_i}[\eta_1(t) - \eta_1(t_1)] \leq a, \quad \inf_{t\in\Delta_k}[\eta_1(t) - \eta_1(t_1)] > a.$$

В этих обозначениях (9) переписывается в виде

$$\sum_{k=2}^{N} P\{B_k\} > 1 - \frac{\varepsilon}{2}. \qquad (9')$$

Учитывая (9′), имеем

$$P\{\sup_{t\in\Delta}[\xi(t) - \xi(t_1)] > 2a - \delta\} \geq P\{\bigcup_{k=1}^{N} B_k A_k\} =$$

$$= \sum_{k=1}^{N} P\{B_k A_k\} = \sum_{k=1}^{N} P\{B_k\} \cdot P\{A_k\} \geq 1-\varepsilon.$$

Так как $\varepsilon > 0$ и $\delta > 0$ можно выбирать сколь угодно малыми, то отсюда следует выполнение (6) для $n=2$. Повторив те же рассуждения, можно было бы, исходя из (6) при $n=2$ доказать(6) для $n=4$ и т.д. Следовательно, для любого отрезка и любого $N > 0$

$$P\{\sup_{t\in\Delta} \xi(t) > N\} = 1$$

Теорема доказана.

§ 2. Достаточные условия для непрерывности. Примеры всюду неограниченных процессов.

Рассмотрим теперь достаточные условия для непрерывности выборочных функций стационарных гауссовских процессов. Наиболее сильный результат, из известных в настоящее время автору, принадлежит Ханту [3] . Достаточные условия, полученные Хантом, сформулированы в терминах спектральных функций. Для переформулировки этих условий в терминах корреляционных функций оказывается полезной следующая

ЛЕММА 3. Если $F(\lambda)$ неубывающая функция ограниченной вариации, а

$$\varphi(h) = \int_0^\infty (1 - \cos\lambda h)\, dF(\lambda),$$

то из того, что для некоторых $C > 0$, $a > 0$ и всех достаточно малых h

$$\varphi(h) \leq \frac{C}{|\ln|h||^a} \qquad (10)$$

следует, что при любом $b < a$

$$\int_0^\infty [\ln(1+\lambda)]^b\, dF(\lambda) < \infty. \qquad (11)$$

Обратно, если (II) выполнено при некотором $b>0$, то (IO) справедливо при любых $a \le b$, $C>0$, для всех достаточно малых h, $|h|<\delta=\delta(C,a)$.

Хант показал, что если спектральная функция $F(\lambda)$ стационарного гауссовского процесса $\xi(t)$ удовлетворяет условию(II) при некотором $b>1$, то почти все выборочные функции процесса $\xi(t)$ непрерывны. С другой стороны, в силу указанной выше леммы 4 из (IO) при $a>1$, следует выполнение (II) для b, $1<b<a$. Поэтому условие (IO) при $a>1$ также достаточное для непрерывности выборочных функций. Таким образом, справедлива следующая

ТЕОРЕМА 2 Для того, чтобы почти все выборочные функции стационарного гауссовского процесса $\xi(t)$ были непрерывны, достаточно, чтобы было выполнено одно из следующих эквивалентных между собой условий:

или при некотором $b>1$

$$\int_0^\infty [\ln(1+\lambda)]^b \, dF(\lambda) < \infty, \tag{I2}$$

или при некоторых $a>1$, $C>0$ для всех достаточно малых h

$$M|\xi(t+h)-\xi(t)|^2 \le \frac{C}{|\ln|h||^a}. \tag{I3}$$

Выведем теперь некоторые достаточные условия для того, чтобы почти все выборочные функции стационарных гауссовских процессов были неограничены на любом конечном интервале.

ТЕОРЕМА 3 Если у стационарного гауссовского процесса $\xi(t)$ существует спектральная плотность $f(\lambda)$ такая, что для некоторых $C>0$, $\lambda_0>0$ и всех $\lambda \geq \lambda_0$

$$f(\lambda) \geq \frac{C}{\lambda(\ln\lambda)^2},$$

то почти все выборочные функции неограничены на любом интервале конечной длины.

Метод доказательства этой теоремы состоит в следующем. Пусть $\eta_n(t)$ взаимно независимые стационарные гауссовские процессы $M\,\eta_n(t)=0$, $n=1,2,\ldots$, такие, что у них существуют спектральные плотности $f_n(\lambda)=\frac{1}{2^{n+1}}$ для $0 \leq \lambda < \frac{1}{2^n}$, $f_n(\lambda)=0$ для $\lambda \geq \frac{1}{2^n}$.

Рассмотрим теперь случайный процесс $\eta(t)=\sum_{n=1}^{\infty} C_n\,\eta_n(t)$, $\sum_{n=1}^{\infty} C_n^2=1$, $\sum_{k=m}^{\infty} C_k^2 \geq \frac{C_1}{m}$ для всех $m \geq m_0$ и некоторого $C_1>0$. Можно проверить, что для этого процесса

$$\lim_{\delta\downarrow 0} P\{\,\omega_\eta(\Delta,\delta) > \sqrt{C_1}\,\} \geq \frac{1}{2}.$$

Из доказанной в § I альтернативы вытекает, что почти все выборочные функции процесса $\eta(t)$ неограничены на любом конечном интервале. Для спектральной плотности $f_\eta(\lambda)$ случайного процесса $\eta(t)$ для некоторых $0<K_1<K_2$ имеют место неравенства

$$\frac{K_1}{\lambda(\ln\lambda)^2} \leq f_\eta(\lambda) \leq \frac{K_2}{\lambda(\ln\lambda)^2}.$$

Таким же свойством неограниченности обладает и стационарный гауссовский процесс $\eta_w(t)$, у которого спектральная плотность $g_w(\lambda)=0$ для $\lambda<w$ и $g_w(\lambda)=f_\eta(\lambda)$ для $\lambda>w$. Замечая теперь, что исходный процесс $\xi(t)$ можно представить в виде суммы двух взаимно независимых стационарных гауссовских процессов, один из которых имеет спектральную плотность вида $\alpha\cdot g_w(\lambda)$, $\alpha>0$, $w>0$, получаем утверждение теоремы.

Примеры неограниченных гауссовских процессов можно также строить, исходя из свойств корреляционных функций. Весьма полезной оказывается следующая лемма, доказанная А.Д.Вентцелем.

ЛЕММА 4. Пусть $\xi_1, \dots, \xi_n$, случайные гауссовские величины $M\xi_i = 0$, $M\xi_i^2 = \sigma_i^2$, $M\xi_i\xi_j = r_{ij} < 0$ $(i \neq j)$ $i,j = 1,\dots,n$. Тогда

$$P\{\max_{1\le i\le n} \xi_i \le a\} \le \prod_{i=1}^{n}\left(\frac{1}{\sqrt{2\pi}\,\sigma_i}\int_{-\infty}^{a} e^{-\frac{x^2}{2\sigma_i^2}}dx\right).$$

Если корреляционная функция стационарного случайного процесса $\xi(t)$ $B(h) = M\xi(t+h)\xi(t)$, выпуклая для $0 \le h \le \delta$, $\delta > 0$, то

$$M[\xi(t_1+h) - \xi(t_1)][\xi(t_2+h) - \xi(t_2)] \le 0, \; |t_2 - t_1| < \delta.$$

Пользуясь леммой 4 можно показать, что если у стационарного гауссовского процесса $\xi(t)$

$$M|\xi(t+h) - \xi(t)|^2 \ge \frac{C}{|\ln|h||}, \quad C > 0, \qquad (I4)$$

для всех достаточно малых h,
а $B(h)$ выпуклая функция, то

$$\lim_{\delta\downarrow 0} P\{\omega_\xi(\Delta,\delta) > \sqrt{C}\} = 1.$$

Применяя далее альтернативу § I получаем, что верна

ТЕОРЕМА 4 Если у стационарного гауссовского процесса $\xi(t)$ корреляционная функция $B(h)$ выпуклая для $h \in (0,\delta)$, $\delta > 0$ и имеет место (I4), то почти все выборочные функции процесса $\xi(t)$ неограничены на любом интервале конечной длины.

§ 3. Условия Гельдера

В случае непрерывных с вероятностью единица стационарных гауссовских процессов встает задача изучения модуля непрерывности выборочных функций. Из общих результатов можно отметить следующий. Если $f(h) \geq 0$, $\frac{h}{f(h)} \to 0$ при $h \to 0$, то вероятность события $\left\{ \overline{\lim_{h \to 0}} \frac{|\xi(t+h) - \xi(t)|}{f(h)} > 1 \right\}$ равна либо нулю, либо единице. Примыкает к этому же кругу вопросов и следующая теорема, являющаяся обобщением результата Бакстера [5], на случай стационарных гауссовских процессов.

ТЕОРЕМА 5 Пусть $\xi_i(t)$ $i = 1,2$ два стационарные гауссовские процесса $M|\xi_i(t+h) - \xi_i(t)|^2 = \varphi_i(h)$ такие, что $\frac{h^2}{\varphi_i(h)} \to 0$, $h \to 0$, а $\lim_{h \to 0} \frac{\varphi_1(h)}{\varphi_2(h)} = \alpha > 1$. Тогда найдется такая последовательность чисел $h_n \downarrow 0$, $n \to \infty$, что с вероятностью единица

$$\lim_{m \to \infty} \frac{1}{m} \sum_{n=1}^{m} \frac{[\xi_1(t+h_n) - \xi_1(t)]^2}{\varphi_1(h_n)} = 1, \text{ а } \lim_{m \to \infty} \frac{1}{m} \sum_{n=1}^{m} \frac{[\xi_2(t+h_n) - \xi_2(t)]^2}{\varphi_2(h_n)} = \frac{1}{\alpha}.$$

Доказательство этой теоремы можно получить, используя тот факт, что для стационарного гауссовского процесса $\zeta(t)$ не дифференцируемого в среднем квадратичном значения

$$\frac{\zeta(t+h_1) - \zeta(t)}{\sqrt{M|\zeta(t+h_1) - \zeta(t)|^2}} \quad \text{и} \quad \frac{\zeta(t+h_2) - \zeta(t)}{\sqrt{M|\zeta(t+h_2) - \zeta(t)|^2}}$$

становятся асимптотически взаимно независимыми случайными величинами, когда $h_1 = const$, а $h_2 \downarrow 0$.

Скажем, что случайный процесс $\xi(t) \in H(\alpha, C)$, $0 < \alpha < 1$, $C > 0$, если для любого $C' > C$ с вероятностью единица равномерно по всем t из каждого интервала Δ конечной длины, существует такое $\delta(\Delta, C')$, что для всех $h < \delta(\Delta, C')$

$$|\xi(t+h) - \xi(t)| \leq C'|h|^{\alpha}.$$

Следующая теорема справедлива для процессов общего типа.

ТЕОРЕМА 6 Для того, чтобы стохастически непрерывный случайный процесс $\xi(t) \in H(\alpha, C)$ достаточно, чтобы

$$\sum_{n=1}^{\infty} \sum_{k=0}^{2^n-1} P\left\{ \left|\xi\left(\tfrac{k+1}{2^n}\right) - \xi\left(\tfrac{k}{2^n}\right)\right| > K_\alpha C \frac{1}{2^{n\alpha}} \right\} < \infty, \quad (15)$$

где $K_\alpha = \frac{2^\alpha - 1}{2^{1+2\alpha}}$. Если при некотором $a > 0$

$M|\xi(t+h) - \xi(t)|^a < \varphi(h) \downarrow 0$ при $h \to 0$, то из

$$\sum_{n=1}^{\infty} n^{\alpha a} \varphi\left(\tfrac{1}{n}\right) < \infty$$

следует, что $\xi(t) \in H(\alpha, C)$ для любого $C > 0$.

Доказательство этой теоремы вполне аналогично доказательству известной теоремы А.Н.Колмогорова о непрерывности выборочных функций [6], см также русский перевод книги Дуба [7] стр. 576.

ТЕОРЕМА 7 Если корреляционная функция $B(h)$ стационарного гауссовского процесса такова, что для всех достаточно малых h

$$M|\xi(t+h) - \xi(t)|^2 \leq C_1 \frac{|h|^{2\alpha}}{|\ln|h||}, \quad (16)$$

то $\xi(t) \in H\left(\alpha, \frac{\sqrt{2C_1}}{K_\alpha}\right)$. Если же для $h \in (0, \delta), \delta > 0$, $B(h)$ выпуклая функция такая, что для всех достаточно малых h и $0 < C_2 < C_1$

$$C_2 \frac{|h|^{2\alpha}}{|\ln|h||} \leq M|\xi(t+h) - \xi(t)|^2 \leq C_1 \frac{|h|^{2\alpha}}{|\ln|h||}, \quad (17)$$

то $\xi(t) \in H(\alpha, \frac{\sqrt{2c}}{K_\alpha})$, но $\xi(t) \bar{\in} H(\alpha, C')$ с вероятностью единица, если $C' < \sqrt{2C_2}$.

Доказательство. Если выполнено (I6), то

$$P\left\{\left|\xi\left(\frac{k+1}{2^n}\right) - \xi\left(\frac{k}{2^n}\right)\right| > CK_\alpha \frac{1}{2^{n\alpha}}\right\} \leq e^{-\frac{C^2 \cdot nK_\alpha^2 \ln 2}{2C_1}}.$$

Поэтому ряд (I5) сходится, когда $C > \frac{\sqrt{2C_1}}{K_\alpha}$. Для доказательства того, что $\xi(t) \bar{\in} H(\alpha, C')$, $C' < \sqrt{2C_2}$, воспользуемся леммой 4. Из этой леммы и (I7) имеем

$$P\left\{\max_{1 \leq k \leq n}\left[\xi\left(\frac{k+1}{n}\right) - \xi\left(\frac{k}{n}\right)\right] \leq \frac{C'}{n^\alpha}\right\} \leq$$

$$\leq \left[P\left\{\left[\xi\left(\frac{k+1}{n}\right) - \xi\left(\frac{k}{n}\right)\right] \leq \frac{C'}{n_\alpha}\right\}\right]^n \leq$$

$$\leq \left[1 - \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left[\frac{C'\sqrt{\ln n}}{\sqrt{C_2}} + 1\right]^2\right\}\right]^n \to 0, \quad n \to \infty,$$

если $\ln n - \frac{1}{2}\left[\frac{C'\sqrt{\ln n}}{\sqrt{C_2}} + 1\right]^2 \to +\infty$, $n \to \infty$. Это будет так, если $C' < \sqrt{2C_2}$. Таким образом $\xi(t) \bar{\in} H(\alpha, C')$ для $C' < \sqrt{2C_2}$. Теорема доказана.

В следующей теореме условия достаточные для того, чтобы $\xi(t) \in H(\alpha, C)$ сформулированы в терминах спектральных функций.

ТЕОРЕМА 8. Если спектральная функция $F(\lambda)$ стационарного гауссовского процесса $\xi(t)$ такая, что

$$\int_0^\infty \lambda^{2\alpha} \ln(1+\lambda)\, dF(\lambda) < \infty, \tag{I8}$$

то $\xi(t) \in H(\alpha, C)$ для любого $C > 0$.

Доказательство. Можно показать, что если выполнено (I8), то для любого $C_1 > 0$ и всех достаточно малых h, $|h| \leq \delta(C_1)$

имеет место (16), и утверждение теоремы 8 является следствием теоремы 7.

Аналогичным путем доказывается следующий результат, впервые полученный другим способом Хантом [3].

ТЕОРЕМА 9 Если спектральная функция $F(\lambda)$ у стационарного гауссовского процесса $\xi(t)$ такова, что

$$\int_0^\infty \lambda^{2\alpha} dF(\lambda) < \infty,$$

то почти все выборочные функции удовлетворяют обобщенному условию Гельдера вида

$$|\xi(t+h) - \xi(t)| < C|h|^\alpha \left(\ln \frac{1}{|h|}\right)^{\frac{1}{2}}$$

равномерно по всем t из любого интервала конечной длины, для любого $C > 0$ и всех достаточно малых h.

В случае стационарных гауссовских процессов, непрерывная дифференцируемость с вероятностью единица эквивалентна условию Липшица. Действительно, если для достаточно малых h, равномерно по t из некоторого интервала конечной длины с вероятностью единица

$$|\xi(t+h) - \xi(t)| < C_\omega |h|,$$

где $C_\omega > 0$ случайная величина, то почти все выборочные функции абсолютно непрерывны. Причем производная, которая также является стационарным гауссовским процессом ограничена с вероятностью единица. Непрерывность производной является следствием теоремы I (§ I).

В заключение отметим, что у гауссовских процессов, корреляционные функции которых аналитические, почти все выборочные функции аналитические, обладающие рядом свойств см. [I], [4].

ЛИТЕРАТУРА

[I] Ю.К.Беляев, Локальные свойства выборочных функций стационарных гауссовских процессов, Теор.вер. и ее примен. т. 5, вып.I (I960)

[2] Р.Л.Добрушин, Свойства непрерывности выборочных функций стационарных гауссовских процессов, Теор.вер. и ее примен. т.5, вып.I (I960).

[3] G.A.Hunt, Random Fourier transforms, Trans. Amer. Math. Soc., v.71 (1951), 38-69. (Есть русский перевод сб.Математика ИИЛ, т.2 № 6(I958), 87-II4).

[4] Ю.К.Беляев, Аналитические случайные процессы, Теор.вер. и ее примен., т.4, вып.4 (I959), 437-444.

[5] G. Baxter, A strong limit theorem for Gaussian processes, Proc. Amer. Math. Soc., v. 7, п. 3, (I956), 522-527.

[6] E.E. Slutsky, Alcuni proposizioni sulla theoria degli funzioni aleatorie, Giorn. Inst. Ital. Attuari, 8 (1937), 183-199. (Есть русский перевод, Труды Ср.Аз.ун-та, сер.матем.(5), 3I, (I939), 3-I5.

[7] Д.Дуб, Вероятностные процессы, ИИЛ, I956.

Математический институт им.В.А.Стеклова, г.Москва.

Aleksandr Yakovlevich Khinchin

by B V Gnedenko

Approved For Release 2009/07/09 : CIA-RDP80T00246A011700340001-4

Александр Яковлевич Хинчин

Дамы и господа!

Сегодня мы собрались здесь, чтобы отдать долг памяти математика, заслуги которого в формировании основных идей и методов современной теории вероятностей признаются всеми.

Одной из характерных особенностей развития современной научной мысли является бурный рост статистических концепций в различных областях естествознания, техники и экономики. С полной определенностью выяснилось, что привлечение методов теории вероятностей к изучению принципиальных проблем физики, биологии, химии, астрономии, а также экономики является не прихотью отдельных исследователей и не преходящей модой, а неизбежностью, вызванной существом дела. В результате теперь обоснованно считают, что законы природы носят статистический характер, обусловленный дискретным строением материи. Известно, что эта точка зрения послужила основой многочисленных успехов во всех областях науки. Само собой разумеется, что указанное обстоятельство должно было сказаться на изменении содержания той ветви математики, которая имеет своей целью изучение случайных явлений. Теория вероятностей не могла оставаться в том состоянии, в каком она находилась в прошлом веке и даже в первые два десятилетия нашего века. Полуинтуитивный подход в определении основных понятий теории вероятностей, характерный для 18 и 19 столетий, не мог удовлетворить не только математиков, но и представителей естествознания. Изолированное положение среди математических наук, которое занимала теория вероятностей еще совсем недавно, находилось в резком противоречии с тем ? ответственным ролью ~~положением~~ какую стала играть ~~занимать~~ она во всей системе наших знаний. В результате одной из самых первоочередных научных задач, сложившихся ко второму десятилетию нашего века, оказалась задача преобразования теории вероятностей в стройную математическую дисциплину с логически отточенными основными понятиями, с широко развитыми специфическими методами исследования, и четко установленными связями с другими ветвями математики. Чтобы теория вероятностей превратилась в действительный метод научного познания природы необходимо было также широко развить её проблематику, глубоко проанализировав особенности постановок математических задач естествознания.

Роль А.Я.Хинчина в решении всего комплекса только что указанных вопросов исключительно велика. И несмотря на разнообразие его научных интересов, они производят впечатление единства и научной целеустремленности. В общих чертах их можно охарактеризовать как систематическое изучение места и значения статистических закономерностей в различных частях математики, и естествознания и техники. Я надеюсь, что дальнейшее подтвердит сказанное.

Александр Яковлевич Хинчин родился 19 июля 1894 г. в селе Кондрово Медынского уезда Калужской губернии, известном и в те времена своей бумажной фабрикой. Его отец по специальности инженер-технолог был главным инженером на указанной фабрике и среди специалистов в области бумагоделательного производства пользовался известностью и авторитетом. Детство, а затем и каникулярные месяцы в период обучения сначала в реальном училище /Москва/, а затем в Московском университете А.Я. проводил в Кондрове. Там он организовал любительский театр, участниками которого стали его сверстники из среды рабочих.

Эти годы ранней юности были одновременно годами увлечения литературой и собственных поэтических проб. Результатом этих увлечений явились несколько томиков стихов, изданных в период с 1912 по 1917 год. Несомненно, что увлечение театром и литературой оказало огромное влияние на формирование Хинчина как одного из самых блестящих лекторов, педагогов и авторов математической литературы. Как в устном, так и в письменном изложении он умел мастерски сочетать превосходную литературную форму с научной глубиной, отчетливостью и ясностью мысли.

С 1911 по 1916 г. Хинчин был студентом физико- математического факультета Московского университета. Там он примкнул к той группе студентов, которая была увлечена идеями теории функций действительного переменного и работала под руководством профессоров Д.Ф.Егорова и Н.Н.Лузина. Первые его самостоятельные научные шаги ~~были~~ относятся как раз к этому времени и были вызваны известными работами А.Данжуа о примитивных функциях. В докладе, прочитанном 6 ноября 1914 года на студенческом математическом кружке, Хинчин предложил естественное обобщение понятия производной. для всего духа идей метрической теории функций. Это понятие прочно вошло в арсенал современной науки под наименованием асимптотической производной. ~~Хинчин Хинчин~~ Под этим наименованием он предложил понимать следующее: если в точке x_0 существует предел

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}, \qquad (1)$$

когда x при стремлении к x_0 пробегает значения, принадлежащие некоторому множеству E, имеющему в точке x_0 плотность 1, то этот предел называется асимптотической производной функции $f(x)$ в точке x_0. В указанном докладе он показал, что так определенное понятие инвариантно относительно выбора множества E. Иными словами, если какие либо два множества E_1 и E_2 имеют в точке x_0 плотность 1, и предел (1) существует как для E_1 так и для E_2, то оба они равны между собой. Понятие асимптотической производной и её использование для целей обобщения понятия интеграла Лебега было предметом первых научных статей Хинчина. Позднее основная идея этого понятия была широко им использована для всестороннего изучения локального поведения измеримых функций.

Говорят, что некоторое свойство осуществляется в данной точке асимптотически, если оно имеет место после удаления множества, имеющего в ней плотность 0. Хинчин предложил называть функцию $f(x)$ асимптотически направленной в точке x, если она становится асимптотически убывающей, возрастающей или постоянной. Функция $f(x)$ асимптотически направлена на данном множестве положительной меры, если она асимптотически направлена почти во всех его точках. Основной результат, выясняющий строение асимптотически направленных функций дается следующей теоремой: чтобы функция $f(x)$ была асимптотически направлена на данном множестве, необходимо и достаточно чтобы её значения в этом множестве с точностью до множества произвольно малой меры совпадали со значениями непрерывной функции, обладающей лишь конечным числом максимумов и минимумов.

Важность понятия асимптотической направленности подчеркивается тем, что функции, обладающие этим свойством, имеют почти всюду на рассматриваемом множестве асимптотическую производную. Условие существования асимптотической производной почти всюду на отрезке было найдено Хинчиным еще в заметке 1917 г. ([2].) Для этого необходимо и достаточно, чтобы данная функция совпадала с непрерывной функцией ограниченой вариации на всем отрезке за исключением быть может множества сколь угодно малой меры. Общая структура измеримых функций выявляется следующим предложением Хинчина: всякая измеримая функция, за исключением, быть может, множества меры нуль, либо имеет асимптотическую производную, либо оба её верхних асимптотических производных числа равны $+\infty$, а оба нижних асимптотических производных числа равны $-\infty$. Вскоре после опубликования работы "Исследования о строении измеримых функций" в журнале Математический сборник, в которой проводился указанный анализ, её полный перевод на французский язык был опубликован редакцией журнала "Fundamenta Mathematica". Это увлечение исследованием глубоких свойств измеримых функций не прошло бесследно ни для математики, ни для выбора последующих направлений работы самим Хинчиным: продолжение развития идей Хинчина, особенно для случая функций многих переменных, осуществляется рядом ученых и в наши дни; дальнейшие работы самого Хинчина как в области теории чисел, так и в теории вероятностей в значительной степени проводились под влиянием его первоначальных интересов.

Для формирования Хинчина как ученого исключительно большое значение имели 1923- 1925 г.г., когда он начал разработку двух направлений математических исследований широкого математического значения. Оба эти направления математической мысли в зачаточной форме имелись уже у Э. Бореля. Одно из них было связано с систематическим изучением метрических свойств различных классов иррациональных чисел, другое - с систематическим использованием средств и понятий теории функций и теории множеств в теории вероятностей.

Чтобы составить представление о характере теоретико-числовых результатов Хинчина, приведем несколько формулировок доказанных им теорем. В работе [27], относящейся к 1926 г., содержится доказательство следующего факта: пусть $\varphi(t)$ положительная функция с монотонно убывающим произведением $t^2\varphi(t)$ Неравенство

$$\left|\alpha - \frac{p}{q}\right| \leq \varphi(q)$$

для почти всех α имеет бесконечное число решений в целых числах p и q тогда и только тогда, когда расходится интеграл $\int_0^\infty t\varphi(t)\,dt$. Ряд законченных изящных результатов Хинчина относится к метрической теории непрерывных дробей. Мы ограничимся здесь формулировкой двух таких теорем [75] и [79]. Пусть $a_1, a_2, \ldots$ – неполные частные разложения иррационального числа α в непрерывную дробь, а $q_1, q_2, \ldots$ – знаменатели подходящих дробей этого разложения. Тогда для почти всех α существуют пределы

$$C = \lim_{n\to\infty} \sqrt[n]{a_1 a_2 \ldots a_n}$$

и

$$\mathcal{D} = \lim_{n\to\infty} \sqrt[n]{q_n}$$

где C и $\mathcal{D}$ абсолютные постоянные /$C = 2,6\ldots$; как позднее нашел П.Леви, $\mathcal{D} = e^{\frac{\pi^2}{12\ln 2}}$/. С точки зрения теории вероятностей эти теоремы можно трактовать как асимптотические свойства для сумм членов последовательностей слабо зависимых величин. К этому же кругу идей относится и известный результат Хинчина под названием закона повторного логарифма. Именно в 1923 г. [10] ему удалось уточнить одну оценку частоты распределения нулей и единиц в двоичном разложении действительных чисел,которая была получена в 1914 г. Харди и Литтльвудом.Если через $\mu(n)$ обозначить уклонение числа единиц,находящихся на первых n местах разложения от $1/2$,то,как они обнаружили,для почти всех чисел $\mu(n)=O(\sqrt{n\ln n})$.В работе [10] удалось доказать,что эту оценку можно заменить на более точную:для почти всех чисел $\mu(n)=O(\sqrt{n\ln\ln n})$.Через год появилась статья [15] ,в которой Хинчин трактовал эту задачу,как задачу теории вероятностей.В терминах теории чисел мы можем сформулировать этот окончательный результат так:для почти всех чисел α имеет место равенство

$$\overline{\lim_{n\to\infty}}\frac{\mu(n)}{\sqrt{2n\ln\ln n}}=1.$$

Это равенство и составляет знаменитый закон повторного логарифма,которому позднее было посвящено большое число превосходных исследований многих ученых.Я хочу сейчас напомнить лишь об одном результате,уточняющем закон повторного логарифма в духе первой из приведенных мной теорем теории чисел.Этот результат был получен уже в начале сороковых годов Эрдешем и Феллером[x)] на основе предшествовавшей работы И.Г.Петровского[xx)],посвященной граничным задачам для уравнения теплопроводности.

---

x/ Erdös P.,On the law of the iterated logarithm,Ann.of Math.,43,419-436,1942

Feller W.,The general form of the so-called law of the iterated logarithm,Trans.Amer.Math.Soc. 54,N 3,373-402,1943

xx/ Petrowsky I.,Zur ersten Randwertaufgabe der Wärmeleitungsgleichung,Compos.Math.1,383-419,1935.

Вопрос можно поставить так: найти все те функции $\varphi(n)$, для которых неравенство

$$\mu(n) < \varphi(n)$$

выполняется для почти всех чисел α при всех n, за исключением быть может конечного их числа. Из результата Хинчина вытекает лишь, что условие

$$\varliminf_{n\to\infty} \frac{\varphi(n)}{\sqrt{2n\log\log n}} > 1$$

достаточно, а условие

$$\varlimsup_{n\to\infty} \frac{\varphi(n)}{\sqrt{2n\ln\ln n}} \geq 1$$

~~[illegible]~~ необходимо. Необходимое и достаточное условие, которому должна удовлетворять функция $\varphi(n)$ состоит в сходимости интеграла

$$\int_A^\infty \frac{1}{t}\,\varphi(t)\, e^{-\frac{\varphi^2(t)}{2}}\,dt\,.$$

Евли говорить об исследованиях Хинчина в области неметрических задач теории чисел, то в первую очередь следует указать на его работы по теории диофантовых приближений и на теорему о сложении последовательностей целых чисел [53], [97], [105]. Эта последняя теорема состоит в следующем: пусть $\{\varphi\}$ – последовательность натуральных чисел, $\varphi(n)$ – число членов этой последовательности, которые не превосходят n. Назовем плотностью последовательности $\{\varphi\}$ и обозначим через $\mathfrak{D}(\varphi)$ нижнюю грань чисел $\frac{\varphi(n)}{n}$. Суммой последовательностей $\{\varphi_1\}, \{\varphi_2\}, \ldots, \{\varphi_k\}$ называется последовательность чисел $\varphi_1 + \varphi_2 + \ldots + \varphi_k$, где каждое φ_i есть либо нуль, либо число последовательности $\{\varphi_i\}$ $(1 \leq i \leq k)$. Хинчин доказал, что если

$$\sum_{i=1}^{k} \mathfrak{D}(\varphi_i) \leq 1,$$

то

$$\mathcal{D}\left(\sum_{i=1}^{k}\varphi_i\right) \geq \sum_{i=1}^{k}\mathcal{D}(\varphi_i)$$

для случая последовательностей с равными плотностями. Опубликование этого результата привлекло внимание многих математиков и вызвало многочисленные попытки распространить его на последовательности в разными плотностями. Долгое время, однако, проблема не поддавалась усилиям. И лишь в 1942 г. Манну, а через год Артину и Шерку, удалось найти полное её решение.

Среди достижений Хинчина в теории диофантовых приближений укажем на принадлежащий ему важный принцип переноса, который связывает решение линейных неравенств в целых числах с диофантовыми приближениями коэффициентов аппроксимирующих линейных форм. Говоря о работах Хинчина в области теории чисел, нельзя не упомянуть превосходные популярные книги, написанные им в различные годы. Среди них я хотел бы особо отметить небольшие книжки [73] и [ ], переведенные на многие языки мира.

Как ни значителен вклад Хинчина в теорию функций и теорию чисел, все же основная его роль в прогрессе математики связана с теорией вероятностей. Оттолкнувшись первоначально от задач связанных с теорией чисел /закон повторного логарифма/ и теории функций /сходимость рядов из независимых случайных величин/, он постепенно включал в орбиту своих интересов все больший и больший круг проблем теории вероятностей. Более того он привлек к разработке её проблем многих молодых московских математиков, положив тем самым начало московской школы в теории вероятностей.

Вслед за работами, посвященными закону повторного логарифма и суммированию рядов со случайными членами, последовали исследования Хинчина, посвященные классическим проблемам суммирования независимых случайных величин. Среди полученных результатов я хотел бы выделить исключительно прозрачное условие применимости закона больших чисел в случае независимых одинаково распределенных слагаемых, которое сводится к существованию конечного математического ожидания [44]. Далее нужно отметить плодотворное понятие относительной устойчивости сумм [74], которое оказалось в самой близкой связи с формулировкой окончательных условий сходимости нормированных сумм независимых слагаемых к нормальному распределению. Для случая одинаково распределенных слагаемых Хинчину удалось одновременно с П. Леви и В. Феллером и независимо от них найти необходимые и достаточные условия сходимости к нормальному закону [79]. Особо нужно отметить работы [47] и [48], которые можно считать началом современной проблематики "больших отклонений". К этому же кругу идей мы должны отнести построение Хинчиным общей теории предельных распределений для сумм независимых случайных величин [91]. Основное предложение развитой им теории может быть сформулировано так: класс предельных распределений для сумм независимых, бесконечно малых случайных величин совпадает с классом безгранично делимых распределений. Доказательство этого факта, а также других предложений теории суммирования потребовало развития и приведения в порядок теории безгранично делимых распределений, незадолго перед тем введенных в рассмотрение Бруно де Финетти и А.Н. Колмогоровым.

Теория суммирования трижды вдохновляла Хинчина на написание монографий. Первая монография на эту тему была им издана в 1927 г. [35], после того как он прочел в Московском университете специальный курс на эту тему. Вторая монография [65] связала классические проблемы суммирования с теорией марковских процессов и исследованиями незадолго до того законченными А.Н.Колмогоровым и И.Г.Петровским. Третья монография [92] давала стройное изложение общих предельных теорем для сумм независимых слагаемых и их применение к классической задаче о сходимости нормированных сумм к нормальному закону. Созданию этой книги также предшествовало чтение специального курса в Московском университете. Этот курс привлек тогда к теории суммирования интересы А.А.Боброва, Д.А.Райкова, а также мои.

В непосредственной связи с теорией суммирования находятся работы Хинчина по арифметике законов распределения, в которых исследуются вопросы представления распределений в виде композиции /произведения/ распределений. Среди полученных Хинчиным результатов отметим следующие: каждое распределение разлагается в произведение безгранично делимого и сходящегося произведения конечной или счетной последовательности неразложимых распределений; деление функций распределения вообще говоря неоднозначно / в силу одного примера Б.В.Гнеденко/. К последнему предложению примыкают работы М.Г.Крейна по продолжению эрмитово-определенных функционалов. Исследования Хинчина по арифметике законов распределения вызвали к жизни ряд работ Д.А.Райкова, Дюге и в последнее время Ю.В.Линника.

Интерес к философским и методологическим вопросам, который Хинчин проявлял еще в период школьного обучения, привел его к ряду ряду публикаций по философским вопросам математики [30], [38], [43], [50], [141]. С другой стороны размышления о роли математики в деле познания закономерностей природы ввели его в проблематику физической статистики [51]. Вопросы статистической физики с этого времени заняли большое место в творчестве Хинчина. И, собственно, все, чем он ни занимался позднее в теории вероятностей так или иначе находило отклик в его исследованиях по статистической физике. Это относится и к его занятиям предельными распределениями для сумм независимых случайных величин, и случайными процессами марковского типа, и к развитой им теории стационарных случайных процессов. В личных беседах он постоянно подчеркивал, что именно размышления над вопросами статистической физики натолкнули его на идею рассмотрения того класса случайных процессов, которые теперь получили наименование стационарных.

Несомненно, что в развитии теории вероятностей за последние пятьдесят лет мысль об изучении случайных процессов является одной из самых плодотворных и оказавших наибольшее влияние не только на структуру всей этой ветви математики, но и на увтановление многочисленных глубоких связей с самыми разнообразными естественно-научными и техническими дисциплинами. Идея рассмотрения стационарных процессов в этом прогрессе теории вероятностей оказалась исключительно важной, а роль Хинчина в оформлении начал теории стационарных случайных процессов исключительно большой. В некоторой мере необходимость рассмотрения такого класса процессов начала уже широко ощущаться в науке. При рассмотрении задач геофизики такие ученые как Тейлор и Келлер приблизились к понятию стационарного процесса. Е.Е.Слуцкий на почве анализа статистических рядов пришел к мысли о том, что определенного типа процессы /являющиеся частным случаем стационарных/ способны в известном смысле иммитировать поведение периодических и почтипериодических процессов. В работе [52], а позднее в [60] Хинчиным был установлен ряд теорем типа закона больших чисел для стационарных последовательностей. В 1933 г. им были даны упрощенные доказательства теорем статистической динамики /также типа закона больших чисел/, ранее доказанных Купманом и Дж. фон Нейманом. Основные его успехи в теории стационарных процессов связаны, однако, с двумя другими его статьями. В первой было дано широкое обобщение известной теоремы Биркхофа, доказанной, с использованием многочисленных специальных допущений, лишь для динамических систем с фазовым пространством в виде конечномерного многообразия. Именно в [58] была дана формулировка и приведено полное доказательство теоремы, которая теперь заслуженно получила название теоремы Биркхофа-Хинчина. В случае дискретного времени она звучит так: если последовательность случайных величин

$$\xi_1, \xi_2, \ldots, \xi_n, \ldots$$

стационарна и математическое ожидание ξ_n конечно, то с вероятностью единица существует конечный предел

$$\lim_{n\to\infty} \frac{1}{n}(\xi_1+\xi_2+\ldots+\xi_n).$$

В статье [67] были заложены основы современной спектральной теории стационарных процессов. Здесь же были даны определения стационарных в широком и узком смыслах случайных процессов. Указанные результаты Хинчина продолжают играть в теории стационарных процессов центральную роль, несмотря на то, что со дня их получения прошло более двадцати пяти лет.

Начиная с 1929 г. Хинчин неоднократно возвращался к рассмотрению проблем статистической физики. Только что говорилось о том, как эти его интересы повлияли на занятия теорией стационарных случайных процессов.С другой стороны его исследования в области предельных теорем для сумм привели к разработке метода получения основных теорем статистической физики. Начиная с 1941 г.,он систематически развивал ту идею,что основные математические задачи статистической физики могут быть сведены к хорошо разработанному аппарату предельных теорем для сумм независимых случайных величин.Особую роль при этом играют локальные предельные теоремы.В значительной степени под влиянием этого убеждения Хинчина в последние 15 лет так оживился интерес к локальным предельным теоремам.Эти идеи широко известны по прекрасным монографиям [113], [132], [136] ,первые две из которых переведены и изданы в ряде стран,в том числе в Германии и США.

Последний период творческий деятельности Хинчина падает на годы 1953-1956.Два типа задач занимали его в это время. С одной стороны это было увлечение новой областью - теорией информации,возникшей в трудах в первую очередь К.Шеннона. Здесь его увлекли вопросы логического упорядочения доказательств,данных творцами тоерии информации без достаточно строгих обоснований.Эти работы Хинчина [143] и [150] благодаря немецкому и английскому переводам,сделались легко доступными для ознакомления.

Второе направление исследований связано с проблематикой теории очередей или,как предложил его называть Хинчин,с теорией массового обслуживания.Интерес к стоящим здесь проблемам возник у него еще в начале тридцатых годов,когда он в результате общественной деятельности оказался тесно связан с работниками московской телефонной сети.Именно к этому времени относятся его исследования [57] и [59], достаточно хорошо известные специалистам. В последних работах его увлекали не отдельные частные задачи обслуживания, а изучение входящего потока требований. Главным образом именно этой задаче посвящена его последняя монография [146] и последние его математические работы [148], [149].

Облик Хинчина как ученого должен быть дополнен еще одной чертой: постоянным интересом к вопросам преподавания как в средней школе, так и в высших учебных заведениях. Он сыграл значительную роль в улучшении учебников по элементарной математике, выступил с рядом статей на методические темы, написал ряд увлекательных брошюр для учителей и превосходных учебников для студентов.

В жизни Хинчин был исключительно требователен к себе. Он много внимания уделял научному воспитанию своих учеников, наводя на темы самостоятельных исследований и поощряя всякое проявление научной самостоятельности и инициативы. Я счастлив, что мне пришлось быть одним из самых близких его учеников и наблюдать его в работе и жизни. Он не признавал недоделанных дел и никогда не позволял себе перекладывать на плечи других работу, за которую он нес ответственность. Он не стремился к внешним почестям и, будучи ученым с мировым именем, членом-корреспондентом Академии наук СССР и академиком Академии Педагогических наук РСФСР, продолжал вести скромный образ жизни, уважая людей за их внутренние качества, а не за занимаемое ими положение.

19 ноября 1959 г. после тяжелой и продолжительной болезни Хинчина не стало. Наша наука понесла крупную утрату. Не стало человека, принявшего на свои плечи значительную долю боль-

шого и нужного труда по строительству новой теории вероятностей. На пути служения науки он приобрел уважение коллег и научной молодежи во всем мире. В тоже время он искренне радовался каждому крупному успеху науки, появлению каждого нового молодого дарования. Я помню как он гордился тем, что в нашей науке появился такой блестящий представитель как В. Дёблин и в тоже время как он переживал преждевременную его гибель от рук [redacted] гитлеровских палачей. Хинчин гармонически сочетал дарования математика классического склада и представителя теоретико-множественной культуры, одновременно он видел в математике мощнейшее вредство изучения [redacted] закономерностей окружающей нас природы. [redacted]

В заключение я хочу поблагодарить собравшихся за внимание и уважение к памяти моего учителя.

Список печатных работ А.Я.Хинчина

1916

1. Sur une extention de l'intêgrale de M.Denjoy; C.R.Acad.Sc. 162,287-290

1917

2. Sur la dêrivation asymptotique,C.R.Ac.Sc.Paris,164,142-145

1918

3. О процессе интегрирования Denjoy,Математ.сбор.,30,1-15

1921

4. Sur la thêorie de L'intêgrale de M.DenjoyИзвестия Иваново-Вознесенского политехнического института 3,49-51

1922

5.Об одном свойстве непрерывных дробей и его арифметических приложениях;Там же 5,27-41

6. К вопросу о представлении числа в виде суммы двух простых чисел;Там же 5,42-48

7. Новое доказательство основной теоремы метрической теории множеств;Там же 6,39-41

1923

8. Sur les suites des fonctions analytiques bornêes dans leur ensemble,Fund.Math.4,72-75

9. Das Stetigkeitsaxiom des Linearkontinuus als Induktionsprinzip betrachtet,Fund.Math.4,164-166

10. Ueber dyadische Brüche,Math.Zeitschr.18,109-116

11. Ein Satz über Kettenbrüche mit aritmetischen Anwendungen, Math.Zeitschr.18,289-306

1924

12. О последовательностях аналитических функций;Матем.сбор.31, 147-151

13. Исследования о строении измеримых функций; Матем. сб. 31, 265-285

14. Sur une théorème général relatif aux probabilités dénombrable, C.R.Ac.Sc.Paris, 178, 617-619

15. Ueber einen Satz der Wahrscheinlichkeitsrechnung, Fund. Math. 6, 9-20

16. Einige Sätze über Kettenbrüche mit Anwendungen auf der diofantischen Approximationen, Math. Ann. 92, 115-125

1925

17. Об одном вопросе теории диофантовых приближений; Изв. Иваново-Вознесенского политех. инст. 8, 32-37

18. Zwei Bemerkungen zu einer Arbeit des Herrn Perron, Math. Zeitschr. 22, 274-284

19. Исследования о строении измеримых функций, гл. 2, Матем. сбор. 32, 377-433

20. Ueber die angenäherte Auflösung linearer Gleichungen in ganzen Zahlen ; Там же 32, 203-219

21. Zur Theorie der diophantischen Approximationen; Там же 32, 277-288

22. Bemerkungen zur metrischen Theorie der Kettenbrüche; там же, 32, 326-329

23. Bemerkungen zu meiner Abhandlung "Ein Satz über Kettenbrüche mit arithmetischen Anwendungen", Math. Zeitschr. 22, 316

24. О петербургской игре, Матем. сбор. 32, 330-341

25. Ueber Konvergenz von Reihen, deren Glieder durch den Zufall bestimmt werden ; там же 32, 668-677 /совместно с А.Н. Колмогоровым/

26. Ueber die Anwendbarkeitsgrenzen des Tschebycheffschen Satzes in der Wachrscheinlichkeitsrechnung, Там же 32, 678-687.

1926

27. Zur metrischen Theorie der diophantischen Approkximationen, Math. Zeitschr. 24, 706-714

28. Ueber das Gesetz der grossen Zahlen, Math. Ann. 96, 152-158

29. Ueber eine Klasse linearer diophantischer Approximationen, Rend. Circ. Mat. Palermo, 50, 170-195

30. Идеи интуиционизма и борьба за предмет в современной математики; Вестник Коммунистической Академии 16, 184-192

1927

31. Recherches sur la structure des fonctions mesurables, Fund. math. 9, 212-279

32. Ueber diophantische Approximationen Höheren Grades, Ma .C . 34, 109-112

33. Диофантовы приближения; Труды Всероссийского Матем. с"езда, 131-137

34. Великая теорема Ферма; Госиздат, 1-52; 2-ое изд. ГТТИ, 1932

35. Основные законы теории вероятностей; Второе изд. ГТТИ 1932

1928

36. Sur la loi forte des grands nombres, C.R Ac.Sc.Paris, 186, 285-287

37. Objection d'une note de M.M.Barzin et Errera, Acad. Royale de Belgique, Bull. class de Sc., 5 sér. 14, 223-224

38. Усиленный закон больших чисел и его значение для математической статистики; Вестник статистики, 29, 123-128

39. Begründung der Normalkorrelation nach der Lindebergschen Methode Известия ассоциации научно-исслед. ин-тов МГУ, 1, 37-45

40. Ueber die Stabilität zweidimensionaler Verteilungsgesetze, Математич. сборник 35, 19-23

41. Ueber die angenäherte Auflösung linearer Gleichungen in ganzen Zahlen, Матем. сборник 35, 31-33

42. Теория чисел;очерк развития за 1917-1927 г.г.;Матем.сборник 35, дополнительный выпуск,1-4

1929

43. Роль и характер индукции в математике;Вестник Коммунист.Акад., 1,5-7.

44. Sur la loi des grands nombres,C.R.Ac.Sc.Paris,188,477-479

45. Sur une généralisation des quelques formules classiques,C.R. Ac.Sc.Paris 188,532-534

46. Ueber ein Problem der Wahrscheinlichkeitsrechnung,Math.Zeit. 29,746-752

47. Ueber die positiven und negativen Abweichungen des arithmetischen Mittels, Math.Ann.101,381-385

48. Ueber einen neuen Grenzwertsatz der Wahrscheinlichkeitsrechnung,Math.Ann.101,745-752

49. Ueber Anwendbarkeitskriterien für das Gesetz der grossen Zahlen, Матем.Сбор.36,78-80

50. Учение Мизеса о вероятностях и принципы физической статистики; Успехи физических наук 9,141-166

1930

51. Die Maxwell - Boltzmannsche Energieverteilung als Grenzwertsatz der Wahrscheinlichkeitsrechnung, Труды семинара по теории вероятностей и матем.статистике при МГУ,1,1-11

1932

52. Sulle successioni stazionarie di eventi,Giorn.Ist.Ital.Attuari Anno 3,3,267-272

53. Zur additieven Zahlentheorie, Матем.сборник 39,27-34

54. Ueber eine Ungleichung, Матем.сбор.39,35-39

55. Sur les class es d'événements équivalents, Матем.сбор.39,40-43

56. Remarques sur les suites d'événements obbéissants à la loi

des grands nombres,Матем.сбор.39,115-119

57. Математическая теория стационарной очереди;Мат.сбор.39,73-84

58. Zu Birkhoffs Lösung des Ergodenproblems,Math.Ann.107,485-488

1933

59. О среднем времени простоя станков,Матем.сбор.40,119-123

60. Ueber stazionäre Reihen züfälliger Variablen,Мат.сбор.40,124-128

61. Ueber ein metrisches Problem der additiven Zahlentheorie, Матем.сбор.40,180-189

62. Zur mathematischen Begründung der statistischen Mechanik, Zetshr.für angew.Math. und Mech.13,101-103

63. Детерминанты Грамма для стационарных рядов,Учёные записки МГУ 1,3-5 /совместно с А.О.Гельфонд/

64. The method of spectral reduction in classical dynamics, Proc.Nat.Acad.Sc,USA 19,567-573

65. Asymptotische Gesetze der Wahrscheinlichkeitsrechnung, Springer 1933;ОНТИ ,1936

1934

66. Zur mathematischen Begründung der Maxwell-Boltzmannschen Energieverteilung,Учёные записки МГУ 2,35-38

67. Korrelationstheorie der stationären stochastischen Prozesse,Math.Ann.109,604-615,Успехи матем.наук вып.5,42-51,1938

68. Eine Verscherfung des Poincaréschen Wiederkehrsatz,Comp.Ma Math.1,177-179

69. Fourierkoeffizienten längs Bahnen im Phasenraum,Мат.сбор41 14-16

70. Случай и как наука с ним справляется,ОНТИ,1-

Союзе;Фронт науки и техники 7,36-46

71. Eine arithmetische Eigenschaft summierbaren Functionen, Матем.сборник 41,11-13

72. Теория вероятностей в дореволюционной России и в Советском

1935

73. Цепные дроби,ОНТИ,2-е изд.ГТТИ 1949,(1-116,) Прага 1952

74. Su una legge dei grand numeri generalizzata,Giornale Ist. Ital.Attuari 6,371-393

75. Metrische Kettenbrüchproblem,Compos.Math. 1,361-382

76. Neuer Beweis und Verallgemeinerung eines Hurvitzschen Satzes, Math.Ann.111,631-637

1936

77. Zur metrischen Kettenbrüche,Compos.Math.3,276-285

78. Ein Satz über lineare diophantische Approximationen,Math.Ann.113,398-415

79. Sul dominio di attrazione della legge di Gauss,Giornale Ist. Ital.Attuari 7,3-18

80. Метрические задачи теории иррациональных чисел,Успехи матем. наук вып.1,7-32

1937.

81. Новый вывод одной формулы П.Леви,Бюллетени МГУ,вып.1,1-5

82. Об арифметике законов распределения,Бюллетени МГУ,вып.1,6-17

83. Об одном признаке для характеристических функций,Бюлл.МГУ,5, 1-3

84. Инвариантные классы законов распределения;Бюлл.МГУ,5,4-5

85. Примеры случайных величин,подчиняющихся устойчивым законам распределения,Бюлл.МГУ,вып.5,6-9

86. Н. Sur les lois stables,C.R.Ac Sc.Paris 202,374- 376 374-376 /совм.с П.Леви/

87. Über die angenäherte Auflösung linearer Gleichungen in ganzen Zahlen,Acta Arithm.2,161-172

88. Ueber singuläre Zahlensysteme,Comp.Math.4,424-431

89. Abschätzungen beim kroneckerschen Approximationssatz,Bull. Ist.Math.Tomsk 1,263-265

90. Ueber Klassenkonvergenz von Verteilungsgesetzen, Bull. Ist. Math. Tomsk 1, 258-262

91. Zur Theorie der unbeschränktteilbaren Verteilungsgesetze, Mat. Сбор. 2(44), 79-119.

1938

92. Предельные распределения для сумм независимых случайных величин, ГОНТИ, 1-116

93. Две теоремы о стохастических процессах с однотипными приращениями, Матем. сбор. 3 45 ,577-584

94. Zur Methode der willkürlichen Funkzionen, там же, 3 45 ,585-589

95. Об унимодальных распределениях, Известия Томского матем. института 2,1-7

96. Теория затухающих спонтанных эффектов; Изв. АН СССР, серия матем. 3,313-332

97. Введение иррациональных чисел; Материалы к совещанию учителей, Наркомпрос РСФСР, 9-12 /перепеч. Математика в школе, № 3, 32-34 за 1939/.

98. Комплексные числа /совм. с П.Я.Дорф/; там же, 39-47

1939

99. О сложении последовательностей натуральных чисел, Матем. сбор. 6 /48/, 161-166

100. О локальном росте однородных стохастических процессов без последействия; Известия АН СССР, 4, серия матем. 487-508

101. Основные понятия математики в средней школе; Математика в школе, № 4, 4-22; № 5, 3-10

102. Всестороннее реальное образование советской молодежи; Математика в школе, № 6, 1-7

1940

103. О преподавании математики; Молодая гвардия, № 9, 142-150

104. Основные математические понятия и определения в средней школе; Учпедгиз, 1-

105. О сложении последовательностей натуральных чисел; Успехи матем. наук, У11, 57-61

1941

106. О математических определениях в средней школе;Математика в школе,№ 1,1-10

107. О понятии отношения двух чисел;Математика в школе № 2, 13-16

108. Об аналитических методах статистической механики;ДАН СССР,33,438-441

109. Средние значения сумматорных функций в статистической механике,ДАН СССР,33,442-445

110. О межмолекулярной корреляции,ДАН СССР,33,487-490

1942

111. Законы распределения сумматорных функций в статистической механике,ДАН СССР,34,61-63

1943

112. Sur un cas de corrélation a posteriori;Матем.сбор.12 /54/, 185-196

113. Математические основания статистической механики;ГТТИ,1-1: 126,USA ,1950

114. Об эргодической проблеме квантовой механики,Изв.АН СССР, 7,167-184

115. Конвексные функции и эволюционные теоремы статистической механики;Изв.АН СССР,серия матем.,7,111-122

116. Восемь лекций по математическому анализу;ГТТИ,3-е изд. 1948;украинское издание 1948,

1946.

117. О задаче Чебышева;Изв.АН СССР,серия матем.10,281-294

118. Элементарное введение в теорию вероятностей /совм.с Б.В. Гнеденко/;1- 128;4-ое изд.1957;Варшава 1952,1954;Бухарест 1953 Прага 1954;Будапешт 1954;Берлин 1955,Пекин 1958,Париж 1960

119. О формализме в преподавании математики;Известия Академии педагогических наук РСФСР,№ 4,7-20

1947.

120. Три жемчужины теории чисел;ГТТИ 1-64,2-е изд.1948;Украинское изд.1949,Берлин 1950,Токио 1956

121. Две теоремы,связанные с ~~[illegible]~~ задачей Чебышева;Изв. АН СССР,серия матем. 11,105-110

122. Об одном предельном случае аппроксимационной теоремы Кронекера;ДАН СССР,56,563-565

123. Об одной общей теореме теории линейных диофантовых приближений;ДАН СССР,56,679-681

1948.

124. Теорема переноса для сингулярных систем линейных уравнений;ДАН СССР,59,217-218

125. К теории линейных диофантовых приближений;ДАН СССР 59, 865-867

126. Принцип Дирихле в теории диофантовых приближений;Успехи математических наук Ш,вып.3,1-28

127. Количественная концепция аппроксимационной теории Кронекера;Изв.АН СССР,серия матем. 12,113-122

128. О некоторых приложениях добавочной переменной;Успехи матем.наук Ш,№ 6,188-200

129. Регулярные системы линейных уравнений и общая задача Чебышева;Изв.АН СССР,серия матем.12,249-258

1949

130. О дробных частях линейной формы;Изв.АН СССР,серия матем. 13,3-8

131. Простейший линейных континуум;Успехи матем.наук,1У,вып.2, 180-197

1950

132. Об аналитическом аппарате физической статистики,Труды Математического ин-та АН СССР,вып.33,1-56

133. Статистическая механика как задача теории вероятностей; Успехи матем.наук,У,вып.3,3-46

134. О суммах положительных случайных величин;ДАН СССР,71,1037-1039

135. Предельные теоремы для сумм положительных случайных величин;Украинск.матем.журнал 2,№ 4,3-17

1951.

136. Математические основания квантовой статистики,ГТТИ,1-256
Berlin 1956

137. О законах распределения чисел заполнения в квантовой статистике;ДАН СССР,78,461-463

138. О некоторых общих теоремах статистической физики;Труды математич.ин-та АН СССР вып.38,345-365

139. Элементы теории чисел;т.1 Энциклопедии элементарной математики,ГТТИ,255-353

1952.

140. О классах эквивалентных событий;ДАН СССР,85,713-714

141. Метод произвольных функций и борьба против идеализма в теории вероятностей;Сборник "Философские вопросы современной физики",изд.АН СССР,522-538;франц.перевод:Questions scientifiques V,Paris,Editions de la nouvelle critique v.1,7-24,1954

142. Советская школа теории вероятностей;Китайский математич. журнал,1-7

1953.

143. Понятие энтропии в теории вероятностей;Успехи матем.наук У111,вып.3,3-20;Бухарест 1955;Берлин 1956,США,1957

144. Краткий курс математического анализа;ГТТИ,1-628;3-е изд. 1957;Бухарест 1955,Пекин 1957

145. Андрей Николаевич Колмогоров /к пятидесятилетию со дня рождения/;Успехи матем.наук У111,вып.3,177-200 /совместно с П.С.Александровым/.

1955

146. Математические методы теории массового обслуживания;Труды матем.ин-та АН СССР,вып.49,1-123;Пекин 1958

147. Симметрические функции на многомерных поверхностях;Сбор. памяти А.А.Андронова;541-574

1856

148. Потоки случайных событий без последействия;Теория вероятностей и её применения,1,вып.1,3-18

149. О пуассоновских потоках случайных событий;Теория вероятностей и её применения,1,вып.3,320-327

150. Об основных теоремах теории информации; Успехи матем.наук,Х1,вып.1,17-75

1959

151. Первое знакомство с теорией вероятностей /совместно с А.М.Яглом/;Детская энциклопедия т.3,211-220;Изд.АН педагогич. наук РСФСР

Mathematical Questions of the Shennov Theory on

Optimal Codification of Information

by R L Dobrushin

Р.Л. Добрушин

МАТЕМАТИЧЕСКИЕ ВОПРОСЫ ШЕННОНОВСКОЙ ТЕОРИИ ОПТИМАЛЬНОГО КОДИРОВАНИЯ ИНФОРМАЦИИ

0.1. Ставший за последнее десятилетие модным термин "теория информации" не имеет до сих пор однозначного толкования. В технической и кибернетической литературе под теорией информации подразумевают обычно всю совокупность приложений математических методов к проблемам задания, переработки, хранения и передачи информации. С точки зрения математики в этих приложениях мало единства, так как они основываются на методах, принадлежащих очень разнообразным разделам математики. При подобном толковании к теории информации нужно отнести всю математическую статистику, многое - из теории случайных процессов, развивающиеся в последние годы исследования мощности ε -сетей в функциональных пространствах (см. [56]), трактуемой как оценка количества информации задаваемого элементом этого пространства), оценки "сложности" алгоритмов математического анализа (см. [3], [92]) и т.п.

Но внутри теории информации в широком смысле этого термина важное место занимает молодая ветвь науки, которую часто (особенно в математической литературе) также называют теорией информации. Мы будем, для определенности, называть ее шенноновской теорией оптимального кодирования информации, так как все, что здесь делается, представляет собой непосредственное развитие идей, содержащихся в замечательной и основоположной работе Клода Шеннона [80], и так как шенноновская теория информации исследует методы передачи информации, основанные на оптимальном выборе методов кодирования и декодирования информации, причем всю специфику этой области создает именно возможность широко варьировать метод кодирования информации. В тех же случаях, когда метод кодирования жестко зафиксирован (так обстоит, например, дело в большинстве статистических проблем, где не во власти статистика менять способ, которым заложена в выборку интересующая его информация и вопрос стоит лишь о выборе оптимального декодирования ("решающего правила"), адекватными задаче оказываются не методы шенноновской теории информации, а более привычные методы современной математической статистики.

0.2. За последние годы выяснилось, что шенноновская теория информации может привлечь внимание математиков не только из-за важности этой теории для приложений, но и просто потому, что в ней возникает много оригинальных математических проблем, интересных и вне связи с техническими применениями теории. Раскрылись неожиданные соприкосновения теории информации с другими математическими науками. Так, при построении кодов полезными оказались методы теории групп (см. разд.4), выяснилась связь общих проблем теории информации и вопросов статистики стохастических процессов (работы М.Пинскера). Работы школы А.Н.Колмогорова показали, что идеи шенноновской теории информации очень полезны и в классических областях математического анализа (теория динамических систем см.[54], [55], теории функции см. [56]). Как это часто бывает в молодых областях математики, некоторые из проблем шенноновской теории информации совсем элементарны по формулировке, но решение их является нетривиальной (а иногда и всерьез трудной) задачей. С другой стороны попытки провести строгое построение основных понятий шенноновской теории информации в достаточно общей обстановке приводят к сложным и довольно громоздким построениям, как всегда в теории вероятностей, основанным на абстрактных методах теории множеств и теории меры. Громоздкость этих построений привела даже некоторых математиквв к сомнениям (см. [24])о том, не разумнее ли отказаться от математизации теории информации. Нам кажется, что необходимость строгого обоснования идей шенноновской теории информации не менее неизбежна, чем, например, аналогичная необходимость в теории стохастических процессов и что даже самые общие конструкции теории информации не сложнее и не абстрактнее, чем, скажем, конструкции современной теории марковских процессов (см. в частности [25]). Подобные сомнения об"ясняются, повидимому тем, что если от возникновения теории цепей Маркова до современного этапа развития теории марковских прецессов прошло около 50 лет, то шенноновская теория информации проделала аналогичный путь всего лишь за 10 лет. Но это быстрое развитие - лишь отражение общих темпов современной науки.

0.3. В этом докладе будет дан обзор с единой точки зрения основных направлений шенноновской теории оптимального кодирования информации. При этом автор, конечно, не надеется достичь равномерности в освещении всех важных вопросов. Ударение будет делаться на вопросах, связанных с работами автора, в том числе и недавними. В то же время некоторые, не менее важные вопросы останутся в тени. Особое внимание будет уделено популяризации очередных ждущих решения и весьма разных по трудности и по конкретности их постановки вопросов шенноновской теории информации. Чтобы выделить их ия текста, не загромождая изложения, мы используем римские цифры на полях.

Приводимая в конце статьи библиография является повидимому довольно полной в отношении статей, опубликованных в математических журналах. Из большой технической литературы отобрано лишь то, что прямо соприкасается с содержанием доклада.

1. Основная проблема Шеннона.

1.0. Мы начнет с того, что оформулируем постановку основной проблемы, решенной К.Шенноном, взяв за основу план изложения, предложенный А.Н.Колмогоровым [53] и развитый в работах автора ([21] , предварительные публикации [17], [20]). Автвру кажется, что концепция Колмогорова является в отличие от предыдущих работ (Файнштейн [31], Хинчин [14], Розенблатт [77], Недома [64], Вольфовиц [93], [94], [96], Блекуелл, Брейман, Томасян [8]) достаточно общей, чтобы охватить все практически интересные случаи, и в то же время она проще и физически естественнее, чем, скажем, концепция, предложенная Перезом [67].

1.1. Передающим устройством (Q,V) 1/ мы будем называть совокупность двух измеримых пространств $(\mathcal{Y}, S_{\mathcal{Y}}), (\tilde{\mathcal{Y}}, S_{\tilde{\mathcal{Y}}})$ и функции $Q(y,A)$, $y \in \mathcal{Y}$, $A \in S_{\tilde{\mathcal{Y}}}$, являющейся измеримой относительно σ-алгебры $S_{\mathcal{Y}}$ при фиксированном $A \in S_{\tilde{\mathcal{Y}}}$ и вероятностной мерой на $(\tilde{\mathcal{Y}}, S_{\tilde{\mathcal{Y}}})$ при фиксированном $y \in \mathcal{Y}$.

1/ Иногда (в т.числе в нашей публикации [17]) тот об"ект, который мы называем передающим устройством, называют каналом. Мы предпочли использовать термин канал для несколько иного об"екта (см.ниже).

и подмножества V в пространстве всех вероятност-
ных мер в произведении $(\mathcal{Y}\times\tilde{\mathcal{Y}}, S_{\mathcal{Y}}\times S_{\tilde{\mathcal{Y}}})$. Пространства
$(\mathcal{Y},S_{\mathcal{Y}})$, $(\tilde{\mathcal{Y}},S_{\tilde{\mathcal{Y}}})$ мы будем называть пространствами сигна-
лов на входе и выходе передающего устройства, и подмно-
жество V – ограничением на распределения сигналов. Две
случайных величины ξ и $\tilde{\xi}$ мы будем называть связанными
передающим устройством (Q,V), если они принимают значе-
ния в $(\mathcal{Y},S_{\mathcal{Y}})$, $(\tilde{\mathcal{Y}},S_{\tilde{\mathcal{Y}}})$, соответственно, для любого
$A\in S_{\tilde{\mathcal{Y}}}$ с вероятностью 1 условная вероятность

$$\mathcal{P}\{\tilde{\xi}\in A/\xi\}=Q(\xi,A) \quad (1)$$

и распределение пары $(\xi,\tilde{\xi})$ принадлежит V.

С наглядной точки зрения $\mathcal{Y}$ –это совокупность переда-
ваемых передатчиком, а $\tilde{\mathcal{Y}}$ – совокупность принимаемых прием-
ником сигналов (в приложениях пространства $\mathcal{Y}$ и $\tilde{\mathcal{Y}}$ часто
совпадают). Если задано случайное значение ξ сигнала на
входе, то (1) позволяет найти условное распределение сигна-
ла на выходе $\tilde{\xi}$. Таким образом $Q(y,\cdot)$ – это распределе-
ние сигнала на выходе, если со входа передан сигнал y.
Наконец, введение ограничения V связано с тем, что во
многих приложениях нельзя считать распределение входного и
выходного сигналов произвольным (типична, например, ситуа-
ция, когда предполагается, что среднее значение квадрата
(мощность) входного сигнала не превосходит заданной кон-
станты). Если, как это делалось в большинстве предыдущих
работ, не хотеть вводить такое ограничение, то в нашем
определении надо взять за V – совокупность всех вероят-
ностных мер на $(\mathcal{Y}\times\tilde{\mathcal{Y}}, S_{\mathcal{Y}}\times S_{\tilde{\mathcal{Y}}})$.

Ограничение V мы будем задавать в дальнейшем не произвольно, а предположим, что имеется $\mathcal{N}$ измеримых функций $\pi_i(y,\tilde{y}),\ y\in Y,\ \tilde{y}\in\tilde{Y}$ и множество $\overline{V}$ в $\mathcal{N}$-мерном евклидовом пространстве, причем распределение $(\eta,\tilde{\eta})$ принадлежит V, лишь если вектор из математических ожиданий

$$\left(E\pi_1(\eta,\tilde{\eta}),E\pi_2(\eta,\tilde{\eta}),\ldots,E\pi_{\mathcal{N}}(\eta,\tilde{\eta})\right)\in\overline{V}. \qquad (2)$$

Предположение о том, что ограничение V задается в виде (2) достаточно для всех практически интересных частных случаев, за исключением, пожалуй, примеров следующего типа. Пусть (Y, S_Y) - пространство действительных функций $y(t),\ t\in[a,b]$ с обычной σ-алгеброй измеримых множеств. Тогда величина η со значениями в пространстве сигналов на входе является случайным процессом $\eta(t),\ t\in[a,b]$. Предположим, что условие V сводится к требованию $E\eta^2(t)\le P_s$, где P_s - заданная константа, для всех $t\in[a,b]$. Подобное ограничение соответствует континууму функций $\pi(\cdot,\cdot)$ и его нельзя задать в виде (2). Остается открытым вопрос о доказательстве для этого и ему аналогичных случаев, формулируемых в разделе 2 теорем Шеннона.

I

Перечислим теперь некоторые важные частные классы передающих устройств.

1.2. Передающее устройство называется <u>отрезком однородного канала без памяти длины</u> n, если имеются измеримые пространства $(Y_0, S_{Y_0}),(\tilde{Y}_0, S_{\tilde{Y}_0})$ и переходная функция $Q_0(y,A),\ y\in Y_0,\ A\in S_{\tilde{Y}_0}$ такие, что пространствами сигналов на входе и выходе передающего устройства являются n-ные степени $(Y_0^n, S_{Y_0}^n),(\tilde{Y}_0^n, S_{\tilde{Y}_0}^n)$ пространства (Y_0, S_{Y_0}), $(\tilde{Y}_0, S_{\tilde{Y}_0})$, т.е. произведения n экземпляров этих пространств) и при любом $y=(y_1,\ldots,y_n)\in Y_0^n$

переходная функция

$$Q(y,\cdot)=Q_0(y_1,\cdot)\times Q_0(y_2,\cdot)\times\ldots\times Q_0(y_n,\cdot),$$

т.е. является декартовым произведением соответствующих мер. Наконец, должно быть задано множество распределений V_0 такое, что распределение величин $\eta=(\eta_1,\ldots,\eta_n)$, $\tilde{\eta}=(\tilde{\eta}_1,\ldots,\tilde{\eta}_n)$ принадлежит V, если и только если пары $(\eta_i,\tilde{\eta}_i)$ имеют распределения, принадлежащие V_0, при всех $i=1,\ldots,n$. Наглядно это определение означает, что i-ая компонента сигнала на выходе $\tilde{\eta}_i$ зависит лишь от i-ой же компоненты η_i сигнала на входе и что ограничение сводится к ограничениям, наложенным на каждую из пар компонент. Если пространства Y_0, $\tilde{Y}_0$ состоят из конечного (или счетного) числа элементов, то переходную функцию $Q_0(\cdot,\cdot)$ можно задавать при помощи матрицы переходных вероятностей (аналогично тому, как это делается в теории марковских процессов). Условие V обычно тогда отсутствует. В этом случае канал без памяти называется <u>конечным</u> (соответственно, <u>счетным</u>).

Наиболее разработанным и важным для приложений частным случаем канала без памяти является <u>канал без памяти с аддитивным шумом</u>. В этом случае (Y_0,S_{Y_0}), $(\tilde{Y}_0,S_{\tilde{Y}_0})$ – действительные прямые, а $Q_0(y,\cdot)$ – совпадает с распределением величины $\zeta+y$, где ζ имеет заданное распределение, причем $E\zeta^2=P_n$ называется <u>мощностью шума.</u> В этом случае обычно в качестве V_0 берут условие $E\eta_i^2\le P_s$, где P_s называют <u>мощностью сигнала.</u> Такой канал называется <u>гауссовским,</u> если ζ имеет гауссовское распределение.

1.3. Значительно более общий класс передающих устройств дает нам понятие <u>однородного канала с дискретным време-</u>

нем. Для задания такого канала надо указать измеримые пространства $(\mathcal{Y}_0, S_{\mathcal{Y}_0})$ и $(\tilde{\mathcal{Y}}_0, S_{\tilde{\mathcal{Y}}_0})$ - сигналов на входе и выходе в любой момент времени, а также измеримое пространство (F, S_F), называемое пространством состояний канала и переходную функцию канала $Q_0(y_0, f, A), y_0 \in \mathcal{Y}_0, f \in F$ $A \in S_{\tilde{\mathcal{Y}}} \times S_F$, измеримую относительно $S_{\mathcal{Y}_0} \times F$ при фиксированном A и являющуюся мерой на $\tilde{\mathcal{Y}}_0 \times F$ при фиксированных y_0, f. С наглядной точки зрения можно представлять себе, что имеются последовательности величин $\ldots, \eta_{-1}, \eta_0, \ldots$ и $\ldots, \tilde{\eta}_{-1}, \tilde{\eta}_0, \tilde{\eta}_1, \ldots$, принимающих значения в пространствах $(\mathcal{Y}_0, S_{\mathcal{Y}_0})$, $(\tilde{\mathcal{Y}}_0, S_{\tilde{\mathcal{Y}}_0})$ соответственно и последовательность величин $\ldots, \varphi_{-1}, \varphi_0, \varphi_1, \ldots$ со значениями в (F, S_F), причем при заданных $\eta_i = y$ и $\varphi_{i-1} = f$ условное распределение для пары $\tilde{\eta}_i, \varphi_i$ задается как $Q_0(y, f, \cdot)$ и не зависит от дальнейших сведений о $\varphi_k, \eta_k, \tilde{\eta}_k$ в предыдущие моменты времени.

Z абзац

Таким образом, состояния канала описывают память канала о прошлом. Точнее, каждому начальному распределению вероятностей $P_0(\cdot)$ на (F, S_F) и целому n сопоставляется передающее устройство, называемое отрезком длины n однородного канала с дискретным временем.

Здесь пространствами сигналов на входе и выходе передающего устройства являются степени $(\mathcal{Y}_0^n, S^n_{\mathcal{Y}_0}), (\tilde{\mathcal{Y}}_0^n, S^n_{\tilde{\mathcal{Y}}_0})$.

Чтобы не загромождать изложения мы построим переходную функцию соответствующего передающего устройства $Q(y, A)$, $y \in \mathcal{Y}_0^n$, $A \in S^n_{\tilde{\mathcal{Y}}_0}$ лишь в том частном случае, когда существуют меры $\mu_{\tilde{\mathcal{Y}}_0}(\tilde{A})$ на $(\tilde{\mathcal{Y}}_0, S_{\tilde{\mathcal{Y}}_0})$ и $\mu_F(B)$ на (F, S_F) такие, что переходная функция

$$Q_0(y_0, f_0, A) = \int_A q_0(y_0, f, \tilde{y}_0, f')\, \mu_{\tilde{\mathcal{Y}}_0}(d\tilde{y}_0)\, \mu_F(df) \qquad (3)$$

т.е. задается плотностью $\dot{q}_0(y_0, f, \tilde{y}_0, f')$. Тогда мы положим

$$Q(y,\tilde{A}) = \int_{\tilde{A}} q(y,\tilde{y})\mu^n_{\tilde{y}_0}(d\tilde{y}),\ y \in Y^n_0,\ \tilde{y} \in \tilde{Y}^n_0,$$

где $y = (y_1,\dots,y_n)$, $\tilde{y} = (\tilde{y}_1,\dots,\tilde{y}_n)$ и

$$q(y_1,\dots,y_n;\tilde{y}_1,\dots,\tilde{y}_n) = \tag{4}$$

$$= \int_F \dots \int_F q_0(y_1,f_0,\tilde{y}_1,f_1)\dots q_0(y_n,f_{n-1},\tilde{y}_n,f_n)\, p_0(df_0)\mu_F(df_1)\dots\mu_F(df_n).$$

С наглядной точки зоения определение (4) можно раз"яснить, сказав, что подынтегральная функция в (4) представляет собой плотность совместного распределения вероятностей для последовтельности состояний канала и последовательности сигналов на выходе канала при заданной последовательности сигналов на входе канала и мы осредняем эту плотность по всевозможным последовательностям состояний канала. Нетрудно дать также более общее определение $Q(\cdot,\cdot)$ для случая, когда $Q_0(\cdot,\cdot)$ не задается плотностью. Далее, по аналогии с тем, как это делается в теории марковских процессов, нетрудно распространить определение отрезка канала на случай неоднородных по времени каналов и каналов с непрерывным временем (Ср.аналогичные построения в [21] § 1.8).

Если пространство F состояний канала конечно, то мы будем говорить о <u>канале с конечным числом состояний.</u> Канал без памяти (см.1.2) можно интрепретировать как канал, в котором пространство состояний состоит из одного элемента

Описанное определение навеяно идеями современной теории автоматов (см. [87]) и его общность достаточна, чтобы охватить большую часть физически интересных случаев. Для каналов с конечной памятью и конечными пространствами сигналов близкое определение изучено в работе Блеквелла, Бреймана и Томасяна (см. [8]) . Аналогичное определение высказывалось также А.Н.Колмагоровым в 1957 году на руководимом им семинаре в Московском Университете.

Отметим, что наше определение не выиграло бы в общности, если бы мы дополнительно предложили, что сигнал на выходе зависит также от сигналов на входе и выходе в предыдущие моменты времени. Дело в том, что от этой зависимости можно всегда избавиться, расширив пространство состояний канала, а именно - назвав состоянием канала совокупность прежнего состояния и значений сигналов на входе и выходе во все предыдущие моменты времени (подробнее см. [8]), где подобным методом показывается, что каналы с конечной памятью в смысле Хинчина-Файнштейна (см. [14], [33]) представляют собой частный случай каналов с конечным числом состояний). Этот прием позволяет также превратить в канал широкий класс передающих устройств, действующих во времени и обладающих свойством отсутствия предвосхищения (соответствующее определение дается по аналогии с определениями в [14], [33]). Однако, при этом может нарушиться свойство однородности по времени, что впрочем не произойдет, если вводить аналогичные понятия не для конечных отрезков канала, а для каналов, действующих бесконечное время.

1.4. Иногда в приложениях оказывается естественным предположить, что канал зависит от медленно меняющихся параметров. В пределе можно предположить, что в течение времени передачи параметры не меняются, но их точное значение неизвестно и является случайным с заданным распределением вероятностей. Тогда мы приходит к понятию канала со случайным параметром. Общее определение этого канала состоит в следующем: предполагается (в обозначениях 1.3), что задано измеримое пространство (B, S_B) -значений параметра, и измеримая функция $\beta(\ell), \ell \in F$ со значениями в (B, S_B) такая, что переходная плотность $q_0(y_0, \ell, \tilde{y}_0, \ell') = 0$ всегда, когда $\beta(\ell) \neq \beta(\ell')$. (Аналогичное определение легко дать и если плотность не существует). Тогда при заданном начальном распределении $P_0(\cdot)$ естественно называть случайным параметром канала случайную величину $\beta(\varphi)$, где φ - имеет распределение $P_0(\cdot)$. С наглядной точки зрения $\beta(\varphi)$ - это не меняющийся во времени случайный параметр. Естественным образом вводится тогда понятие условного канала при заданном $b \in B$. Пространством состояний этого канала является прообраз точки b при отображении $\beta(\ell), \ell \in F$, а переходная плотностью совпадает с $q_0(y_0, \ell, \tilde{y}_0, \ell')$, где $\beta(\ell) = \beta(\ell') = b$.

1.5. Передающее устройство называется гауссовским, если пространство сигналов на входе и на выходе являются пространствами функций с действительными значениями, переходная функция $Q(y, \cdot)$ задает при любом фиксированном y - условное гауссовское распределение (мы называем бесконечную совокупность случайных величин, имеющей условное гауссовское распределение, если любая конечная подсовокупность имеет при любом условии конечномерное гауссовское распределение со вторыми моментами, не зависящими от условия, и с первыми моментами, линейно зависящими от этого условия) и, наконец, ограничение V наложено лишь на первые и вторые моменты, связанных передающим устройством случайных величин. Такое передающее устройство ставит в соответствие гауссовской величине η на входе, гауссовскую же величину $\tilde{\eta}$ на выходе.

. В приложениях особо важную роль играют приемно-передающие устройства с конечной полосой пропускания. При попытке сформулировать математически точно соответствующее понятие и доказать известную формулу Шеннона [35], рассматривая канал, действующий бесконечное время, возникают серьезные трудности, связанные с тем, что, как известно, стационарные процессы с ограниченным спектром сингулярны (детерминированы) и поэтому не несут информации. Обойти эти трудности можно, введя следующую модель передающего устройства с конечной полосой пропускания $[\lambda, \lambda+W]$ и аддитивным шумом для передачи за время T, которая, быть может, несколько кустарна, с точки зрения общей теории стохастических процессов, но достаточно хорошо отражает реальную физическую ситуацию (ср. [43]).

Здесь пространства (Y, S_Y) и $(\tilde{Y}, S_{\tilde{Y}})$ – пространства функций $y(t)$, $t \in [0,T]$ с обычным образом вводимой σ-алгеброй измеримых множеств. Далее задан случайный процесс $\zeta(t), t \in [0,T]$ (шум.) Оператор A сопоставляет функции $v(t), t \in [0,T]$ функцию

$$Av(t) = \sum_{\lambda_0 \le k/T \le \lambda_0 + W} \left(c_k \cos\frac{2\pi k t}{T} + d_k \sin\frac{2\pi k t}{T}\right),$$

где c_k, d_k – коэффициенты Фурье функции $v(t)$. Тогда мера $Q(y(t), \cdot)$ представляет собой распределение случайного процесса $A(y(t)+\zeta(t))$. Условие V -состоит в ограниченности средней мощности сигнала на входе, т.е. оно состоит в том, чтобы $E\left\{\int_0^T \eta^2(t)dt\right\} \le P_s T$ где P_s - заданная константа, называемая средней мощностью сигнала. Средней мощностью шума P_n назовем $\frac{1}{T}E\left\{\int_0^T \zeta^2(t)dt\right\}$. Мы будем говорить, что шум является белым гауссовским, если

$$\zeta(t)=\sum_{\lambda_0 \le \frac{k}{T} \le \lambda_0+W}\left(c_k \cos\frac{2\pi kt}{T}+d_k \sin\frac{2\pi kt}{T}\right),$$

где c_k и d_k - независимые гауссовские величины с нулевыми средними и равными дисперсиями.

1.6. Сообщением (p,W) мы будем называть совокупность двух измеримых пространств (X,S_X) и $(\tilde{X},S_{\tilde{X}})$, распределения вероятностей $P(\cdot)$ в пространстве (X,S_X) и подмножества W в множестве всех вероятностных мер в произведении $(X\times\tilde{X}, S_X\times S_{\tilde{X}})$. Пространства (X,S_X) и $(\tilde{X},S_{\tilde{X}})$ мы будем называть пространствами значений сообщения на входе и выходе, распределение $P(\cdot)$ - входным распределением сообщения и подмножество W -условием точности воспроизведения. Две случайные величины ξ и $\tilde{\xi}$ мы будем называть образующими сообщение (p,W), если они принимают значение в пространствах (X,S_X), $(\tilde{X},S_{\tilde{X}})$, соответственно, распределение величины ξ совпадает с $P(\cdot)$ и совместное распределение пары принадлежит W.

Таким образом, мы, как это обычно делается в теории информации, считаем подлежащее передаче сообщение случайным с заданным распределением $P(\cdot)$. При этом мы считаем, что сообщение, получаемое после передачи, не должно совпадать точно с входным сообщением (в соответствии с нашим определением, вообще говоря, даже не должны совпадать пространства значений сообщений на входе и выходе; хотя в большинстве приложений они все же совпадают). Однако указаны некоторые границы, в которых может изменяться полученное сообщение в зависимости от переданного сообщения. Это и есть условия точности W .

Мы будем предполагать, что условие W задано при помощи M функций $\rho_i(x,\tilde{x})$, $x \in X$, $\tilde{x} \in \tilde{X}$ и множества $\widetilde{W}$ аналогично тому, как (2) задает ограничение V . По поводу его общености можно сделать те же замечания, что и по поводу общности условия (2). В частности остается открытым вопрос,

II аналогичный I.

1.7. Важнейшим частным случаем сообщения является <u>сообщение с условием полного воспроизведения</u>, как мы будем называть сообщение, в котором совпадают пространства (X, S_X) и $(\tilde{X}, S_{\tilde{X}})$, а W - такового, что пара $(\xi, \tilde{\xi})$ образует сообщение (p, W) , если и только если $\xi = \tilde{\xi}$ с вероятностью 1.

Именно к этому случаю сводится при нашей трактовке теорема Шеннона исследования, проводившееся в большой серии работ, начатой Файнштейном [31] и Хинчиным [14]. Более общая трактовка сообщения, введенная еще Шенноном [80], математически уточнена Колмогоровым [53] и развита в работах автора. В недавней работе Шеннона [85] независимо исследован частный класс сообщений, называемых в нашей терминологии сообщениями с покомпонентным условием точности воспроизведения.

1.8. Предположим, что заданы измеримые пространства (X_0, S_{X_0}), $(\tilde{X}_0, S_{\tilde{X}_0})$ такие, что $(X, S_X) = (X_0^n, S_{X_0}^n)$, $(\tilde{X}, S_{\tilde{X}}) =$

$= (\tilde{X}_0^n, S_{\tilde{X}_0}^n)$. Тогда случайные величины ξ и $\tilde{\xi}$, принимающие значение в X и $\tilde{X}$ можно трактовать как наборы $\xi = (\xi_1, \ldots, \xi_n)$, $\tilde{\xi} = (\tilde{\xi}_1, \ldots, \tilde{\xi}_n)$, где ξ_i принимают значения из X_0, а $\tilde{\xi}_i$ из $\tilde{X}_0$. Величины ξ_i мы будем называть тогда компонентами сообщения на входе, а $\tilde{\xi}_i$ - компонентами сообщения на выходе. Мы будем говорить, что условие точности воспроизведения W является покомпонентным, если выполнение или нев... ...ние этого условия для пары $\xi, \tilde{\xi}$ зависит лишь от попарных распределений пар $(\xi_i, \tilde{\xi}_i), i=1,\ldots,n$, но не от совместных распределений этих пар. Естественным образом вводится понятие однородного по времени сообщения. Покомпонентное условие точности воспроизведения будем называть однородным условием ограниченности среднеквадратичной ошибки, если (X_0, S_{X_0}), $(\tilde{X}_0, S_{\tilde{X}_0})$ - действительные прямые и условие состоит в том, чтобы $E(\xi_i - \tilde{\xi}_i)^2 \le P_d$, где P_d - заданная константа, называемая среднеквадратичной ошибкой. Мы будем говорить, что условие точности воспроизведения W является аддитивным, если каждаяи из функций $\rho_i(x, \tilde{x})$ имеет вид

$$\rho_i(x,\tilde{x}) = \sum_{i=1}^{n} \rho_i^0(x_i^0, \tilde{x}_i^0), \text{ где } x = (x_1^0, \ldots, x_n^0), \tilde{x} = (\tilde{x}_1^0, \ldots, \tilde{x}_n^0).$$

Мы будем говорить, что сообщение на входе имеет независимые компоненты, если для величины ξ с распределением $P(.)$ компоненты $\xi_1, \ldots, \xi_n$ являются независимыми величинами.

1.9. Другим важным частным случаем является гауссовское сообщение. Здесь предполагается, что пространства значений сообщений на входе и на выходе являются пространствами функций с действительными значениями, так что величины ξ и $\tilde{\xi}$ можно трактовать как наборы действительных величин.

Сообщение называется гауссовским, если распределение $p(\cdot)$ задает набор величин, имеющих совместные гауссовские распределения, а условие W наложено лишь на первые и вторые моменты рассматриваемых случайных величин.

1.10. Пусть теперь одновременно заданы передающие устройства (Q,V) и сообщение (p,W). Мы будем называть кодированием функцию $\mathcal{P}(x,A), x\in X, A\in S_{\mathcal{Y}}$, являющуюся при фиксированном значении x сообщения на входе вероятностной мерой на пространстве $\mathcal{Y}$ сигналов на входе, а при фиксированном A измеримую относительно S_X. Мы будем называть декодированием функцию $\widetilde{\mathcal{P}}(\tilde{y},\tilde{A}), \tilde{y}\in\tilde{\mathcal{Y}}, \tilde{A}\in S_{\tilde{X}}$, являющуюся при фиксированном сигнале на выходе $\tilde{y}$ мерой на пространстве $\tilde{X}$ сообщений на выходе, а при фиксированном $\tilde{A}$ измеримой относительно $S_{\tilde{\mathcal{Y}}}$. Мы будем говорить, что сообщение (p,W) может быть передано через передающее устройство (Q,V) при помощи кодирования $\mathcal{P}(\cdot,\cdot)$ и декодирования $\widetilde{\mathcal{P}}(\cdot,\cdot)$, если могут быть построены случайные величины $\xi, \eta, \tilde{\eta}, \tilde{\xi}$, образующие цепь Маркова такие, что пара $\xi, \tilde{\xi}$ образует сообщение (p,W), пара $(\eta,\tilde{\eta})$ связаны передающим устройством (Q,V) и с вероятностью 1

$$\mathcal{P}\{\eta\in A/\xi\} = \mathcal{P}(\xi,A), \quad \mathcal{P}\{\tilde{\xi}\in\tilde{A}/\tilde{\eta}\} = \widetilde{\mathcal{P}}(\tilde{\eta},\tilde{A}). \qquad (5)$$

Мы будем говорить в этом случае, что величины $\xi, \eta, \tilde{\eta}, \tilde{\xi}$ задают способ передачи сообщения (p,W) через передающее устройство (Q,V). ~~Хххх~~ Заметим, что задание кодирования $\mathcal{P}$ и декодирования $\widetilde{\mathcal{P}}$ однозначно определяет совместное распределение величин $\xi, \eta, \tilde{\eta}, \tilde{\xi}$.

Прокомментируем это определение. Использование кодирования $\mathcal{P}(\cdot,\cdot)$ означает с наглядной точки зрения, что если сообщение приняло значение x, то мы передаем сигнал на входе, выбранный с распределением $\mathcal{P}(x,\cdot)$. В большинстве предыдущих работ использовалось лишь <u>нерандомизированное кодирование</u>, задаваемое функцией $f(x)$, $x \in X$ со значениями в Y такой, что мера $\mathcal{P}(x,\cdot)$ сосредоточена в точке $f(x)$ и, аналогично, определяемое <u>нерандомизированное декодирование</u>. Конечно, на практике почти всегда используют нерандомизированные кодирование и декодирование, но введение рандомизации заведомо входит в принципиальные возможности создателя системы связи. С математической точки зрения оно упрощает формулировки теорем. Нетрудно доказать, что если ограничение V отсутствует, в определение условия W входит лишь одна функция $\rho_1(\cdot,\cdot)$, $(M=1)$ и $\overline{W}=[0,a]$, то из <u>возможности как либо передать сообщение</u> (p,W) <u>через передающее устройство</u> (Q,V) <u>вытекает возможность такой передачи при помощи нерандомизированных кодирования и декодирования.</u> Нетрудно привести примеры, показывающие, что если $M=2$, это утверждение может быть не верно. Однако все же можно ожидать, что при очень широких условиях (быть может в

III некотором асимптотическом смысле (ср. разд. 2.1b)) верна некоторая теорема о возможности замены рандомизированных кодирования и декодирования на нерандомизированные.

Предположение о том, что $\xi, \eta, \tilde{\eta}, \tilde{\xi}$ образуют цепь Маркова является совсем естественным. Оно означает наглядно, что при передаче сигнал на входе зависит лишь от сигнала на входе, а не от того, какое значение сообщения им закодировано, и что при декодировании используется лишь сигнал на выходе, а не недоступные сигнал и сообщение на входе.

1.11. Мы можем сформулировать теперь основную проблему Шеннона. Для каких пар сообщение (p, W) и передающее устройство (Q, V) возможно и для каких невозможно подобрать кодирование и декодирование так, чтобы (p, W) можно было передать через передающее устройство (Q, V). Решения этой проблемы при разных предположениях мы будем называть теоремами Шеннона.

Естественно возникает также вторая проблема о том, как в случае, когда передача возможна, построить по возможности простым и эффективным образом кодирование и декодирование, эту передачу осуществляющие. Исследование этих двух проблем, а также их непосредственных обобщений, составляет в настоящее время предмет Шенноновской теории оптимального кодирования информации.

2. Теоремы Шеннона.

2.1. В основополагающей работе Шеннона были введены величины, позволяющие сформулировать ответ на поставленную им проблему. Главной из них является величина, называемая информацией.

Пусть заданы две случайные величины ζ и $\tilde{\zeta}$, принимающие значения в измеримых пространствах (Z, S_Z) и $(\tilde{Z}, S_{\tilde{Z}})$, соответственно. Пусть $P_{\zeta\tilde{\zeta}}(\mathcal{D}), \mathcal{D} \in S_Z \times S_{\tilde{Z}}$, $P_\zeta(C), C \in S_Z, P_{\tilde{\zeta}}(\tilde{C}), \tilde{C} \in S_{\tilde{Z}}$ их совместное и одномерные распределения. Если мера $P_{\zeta\tilde{\zeta}}(\cdot)$ не является абсолютно непрерывной относительно произведения мер $P_\zeta \times P_{\tilde{\zeta}}(\cdot)$, то мы положим информацию

$$J(\zeta, \tilde{\zeta}) = +\infty. \tag{6}$$

Если же $P_{\zeta\tilde{\zeta}}(\cdot)$ абсолютно непрерывно относительно

$P_\xi \times P_{\tilde\xi}(\cdot)$, то мы обозначим через $a_{\xi\tilde\xi}(\cdot)$ плотность по Радону-Никодиму меры $P_{\xi\tilde\xi}(\cdot)$ относительно $P_\xi \times P_{\tilde\xi}(\cdot)$. и назовем информационной плотностью функцию

$$i_{\xi\tilde\xi}(\cdot) = \log a_{\xi\tilde\xi}(\cdot) \qquad (7)$$

Информацией мы будем называть

$$J(\xi,\tilde\xi) = \int\limits_{Z\times\tilde Z} i_{\xi\tilde\xi}(z,\tilde z) P_{\xi\tilde\xi}(dz,d\tilde z) = \int\limits_{Z\times\tilde Z} a_{\xi\tilde\xi}(z,\tilde z)\log a_{\xi\tilde\xi}(z,\tilde z) P_\xi(dz) P_{\tilde\xi}(d\tilde z). \qquad (8)$$

Если (Z, S_Z) и $(\tilde Z, S_{\tilde Z})$ – действительные прямые, а распределения $P_{\xi\tilde\xi}(\cdot), P_\xi(\cdot), P_{\tilde\xi}(\cdot)$ задаются плотностями $q_{\xi\tilde\xi}(z,\tilde z), q_\xi(z), q_{\tilde\xi}(\tilde z)$, то определение (8) принимает более классическую форму

$$J(\xi,\tilde\xi) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \log \frac{q_{\xi\tilde\xi}(z,\tilde z)}{q_\xi(z) q_{\tilde\xi}(\tilde z)} dz\, d\tilde z. \qquad (9)$$

Энтропией случайной величины ξ называется ее информация относительно самой себя

$$J(\xi,\xi) = H(\xi). \qquad (10)$$

Часто в качестве основного определения информации берут вместо (8) выражения информации как предела (или верхней грани) некоторой последовательности интегральных сумм. Соотношение между различными возможными определениями информации подробно исследовано в серии работ Переза [36], Гельфанда, Колмогорова, Яглома [38] [40], Цзян-Цзе-пея [100] Добрушина [21] 1/.

1/ Несколько запоздавшая работа ← [36] является отражением недостаточности контакта между советскими и американскими исследователями в этой области.

2.2. Основным доказательством научной важности понятий информации и энтропии кажется нам то, что при их помощи дается решение проблемы Шеннона. Имеются отдельные и пока что разрозненные указания на то, что понятия энтропии и информации могут оказаться ценными и вне обстановки, описанной в разделе 1 (см. раздел 7), однако в целом надо сказать, что многочисленные попытки использовать математические понятия энтропии и информации всюду, где применимо слово информация в общелексической трактовке этого слова, лишь компрометирует эти понятия. С этой точки зрения очень популярное аксиоматическое определение энтропии, введенное еще Шенноном [80] и затем упрощенное другими авторами (Хинчин [13], Фаддеев [30],) сыграло даже отрицательную роль. Дело в том, что не все аксиомы, входящие в определение энтропии, одинаково естественны (например, формула условной энтропии), а также в том, что как отмечал А.Н.Колмогоров, априори не ясно, существует ли единая одномерная численная характеристика, адекватно отражающая свойства такого сложного явления, каким является информация в многообразной, реальной обстановке. Тем не менее кажется интересным дать также аксиоматическое определение информации.
IV
Повидимому основными аксиомами должны быть здесь свойства неотрицательности информации (11), формула условной информации (12) и, быть может, формула, связывающая информацию и энтропию (10).

2.3. Отметим некоторые важные свойства информации.

Информация

$I(\xi,\tilde{\xi}) \geq 0$, причем $I(\xi,\tilde{\xi})=0$ (11)

тогда и только тогда, когда ξ и $\tilde{\xi}$ независимы. Ее зависимость от распределения $P_{\zeta\tilde{\zeta}}(\cdot)$ не является непрерывной, а лишь полунепрерывной снизу (Гельфанд и Яглом [40]). Об"ясняется это тем (см. [23]), что от распределения (с метрикой сходимости по вариации) непрерывно зависит информационная плотность (в смысле сходимости по вероятности), а из сходимости по вероятности последовательности подынтегральных функций не следует сходимость последовательности их интегралов, если эти интегралы не являются равномерно сходящимися. Если существует условное распределение вероятностей для пары ζ и $\tilde{\zeta}$ при условии, что задана третья величина γ , то естественным образом, можно определить условную информацию $I(\zeta,\tilde{\zeta}/\gamma)$ 1/, являющуюся, конечно, случайной величиной. Тогда верна следующая важная <u>формула условной информации</u>

$$I((\zeta,\gamma),\tilde{\zeta}) = I(\gamma,\tilde{\zeta}) + E\,I(\zeta,\tilde{\zeta}/\gamma). \tag{12}$$

Из (11) и (12) легко следует, что если $\varphi(\cdot)$ - некоторая измеримая функция, то

$$I(\varphi(\zeta),\tilde{\zeta}) \le I(\zeta,\tilde{\zeta}). \tag{13}$$

Далее из (12) следует, что если величины $\zeta,\gamma,\tilde{\zeta}$ образуют цепь Маркова, то

$$I((\zeta,\gamma),\tilde{\zeta}) = I(\gamma,\tilde{\zeta}). \tag{14}$$

2.4. Мы можем применить теперь полученные результаты к проблеме Шеннона. Пусть величины $\xi,\eta,\tilde{\eta},\tilde{\xi}$ задают

1/ Можно дать определение условной информации и без предположения о существовании условных распределений (см. [21], [54]).

способ передачи. Тогда, применяя последовательно (13) и (14), находим, что

$$\mathcal{J}(\xi,\tilde{\xi}) \le \mathcal{J}((\xi,\eta),(\tilde{\eta},\tilde{\xi})) \le \mathcal{J}(\eta,\tilde{\eta}), \tag{15}$$

т.е. информация между сообщениями на входе и выходе не может быть больше информации между сигналами на входе и выходе.

Назовем теперь пропускной способностью передающего устройства (Q,V) число

$$C(Q,V) = \sup \mathcal{J}(\eta,\tilde{\eta}), \tag{16}$$

где верхняя грань берется по всем парам величин η, $\tilde{\eta}$, связанным передающим устройством (Q,V). Назовем W -энтропией сообщения число

$$H(p,W) = \inf \mathcal{J}(\xi,\tilde{\xi}), \tag{17}$$

где нижняя грань берется по всем парам $\xi,\tilde{\xi}$, образующим сообщение (p,W). (Название энтропия оправдывается тем, что для сообщения с условием полного воспроизведения $H(p,W) = H(\xi)$). Тогда из (15) вытекает следующая теорема Шеннона: Если сообщение (p,W) может быть передано через передающее устройство (Q,V), то

$$H(p,W) \le C(Q,V). \tag{18}$$

мы будем называть ее обратной теоремой Шеннона.

Это общее и простое доказательство утверждения обратной теоремы Шеннона, основанное лишь на "алгебраических" свойствах информации, было повидимому впервые указано Колмогоровым [53]. Кажется, что оно еще недостаточно популярно в литературе по теории информации, поскольку до сих пор в некоторых работах факты, являющиеся частным случаем этого результата, доказываются более громоздкими, прямыми методами.

2.5. С утверждением теоремы Шеннона о достаточных условиях для возможности передачи дело обстоит сложнее. Нетрудно привести примеры, показывающие, что условие (18) (даже если заменить в нем символ $\leq$ на $<$) недостаточно для возможности передачи. Условие (18) является достаточным для возможности передачи лишь в некотором асимптотическом смысле и при этом главным является предположение о том, что $H(p,W) \to \infty$, так что условие (18) является необходимым и достаточным лишь применительно к задаче о передаче достаточно большого количества информации. Остальные многочисленные предположения носят характер "предположений регулярности" и их можно смело считать выполненными в прикладных исследованиях.

Наиболее уязвимым с точки зрения практической приложимости является в теореме Шеннона то, что методы кодирования и декодирования, существование которых гарантируется выполнением условия (18), повидимому неизбежно связаны с кодированием большого количества информации, как единого целого, и поэтому неизбежно сложны (причем сложность их возрастает при $H/C \to 1$). В связи с этим очень заманчивой, хотя пока что и отдаленной кажется перспектива получения количественных оценок минимальной сложности алгоритмов оптимального кодирования и декодирования (ср. работу [3]), где подобные оценки получаются для алгоритмов вычислительного анализа). С этой точки зрения очень интересны работы [98].

и [28], где делаются оценки сверху для среднего числа действий в алгоритмах декодирования для простейших ситуаций.

2.6. Начиная вводить понятия, используемые в математической формулировке прямого утверждения теоремы Шеннона, мы назовем последовательность пар случайных величин $(\xi^t, \tilde{\xi}^t)$, $t = 1, 2, \dots$, информационно-устойчивой, если $0 < J(\xi^t, \tilde{\xi}^t) < \infty$ и

$$\lim_{t\to\infty} \frac{i_{\xi^t,\tilde{\xi}^t}(\xi^t, \tilde{\xi}^t)}{J(\xi^t, \tilde{\xi}^t)} = 1 \tag{19}$$

в смысле сходимости по вероятности.x/

Во многих приложениях ξ^t и $\tilde{\xi}^t$ естественно считать последовательностями отрезков $[0,t]$ некоторых процессов $\{\alpha_s\}$ и $\{\tilde{\alpha}_s\}$ соответственно (т.е. $\xi^t = \{\alpha_s, s\in[0,t]\}$, $\tilde{\xi}^t = \{\tilde{\alpha}_s, s\in[0,t]\}$). Если процессы α_s и $\tilde{\alpha}_s$ имеют дискретное время и пары $(\alpha_s, \tilde{\alpha}_s)$, $s = \dots, -1, 0, 1, \dots$ ~~взаимно~~ в совокупности независимы, то утверждение об информационной устойчивости отрезков $\xi^t, \tilde{\xi}^t$ сводится к утверждению об законе больших чисел для последовательности сумм независимых случайных величин, и поэтому необходимые и достаточные условия для информационной устойчивости нетрудно получить из теории предельных теорем для сумм независимых величин. Аналогично, если пары $(\alpha_s, \tilde{\alpha}_s)$ образуют цепь Маркова, то утверждение об информационной устойчивости сводится к утверждению о законе больших чисел для функций от величин, связанных в цепь Маркова. Эти замечания

----

x/ Мы не рассматриваем здесь случай бесконечной информации, для которого тоже возможна содержательная теория (см. [21]).

были сделаны впервые Розенблат-Ротом [76] . Если α_s , $\tilde{\alpha}_s$ -образуют совместно стационарный и эргодический процесс с дискретными временем и пространством состояний, то утверждение об информационной устойчивости легко получить из известной теоремы Мак-Миллана [60], а сама эта теорема означает информационную устойчивость последовательности пар $\{(\zeta^t, \zeta^t)\}$. Для стационарных процессов с произвольным множеством состояний и процессов с непрерывным временем первые результаты об условиях информационной устойчивости были получены Перезом ([67], [69]), а в достаточно широких условиях они установлены Пинскером. Здесь очередной задачей кажется обобщение этих результатов на случай "схемы серий" (ср. [41]), т.е. на случай, когда $\zeta^t = \{\alpha_s^t, s \in [0,t]\}$, $\tilde{\zeta}^t = \{\tilde{\alpha}_s^t, s \in [0,t]\}$. Недавно Якобс [46] ввел интересный класс почти-периодических процессов и доказал для них теорему Мак-Миллана, что позволяет установить и их информационную устойчивость. Кажется актуальным перенести результаты Якобса на почти-периодические процессы с непрерывными временем и пространством состояний, поскольку именно такова статистическая природа многих из процессов, возникающих в задачах радиотехники в связи с модуляциями стационарных процессов.

VI

VII

Интересным и окончательным является результат Пинскера, показавшего, что если $\zeta^t, \tilde{\zeta}^t$ представляют собой наборы величин, имеющих совместные гауссовские распределения, то для информационной устойчивости необходимо и достаточно, чтобы $\mathcal{I}(\zeta^t, \tilde{\zeta}^t) \to \infty$.

Из сказанного ясно, что свойство информационной устойчивости верно в очень широкой обстановке. Посвоей всеобщности информационную устойчивость можно сравнить с родственным ей законом больших чисел.

2.7. Рассмотрим далее последовательность передающих устройств (Q^t, V^t) с $C(Q^t, V^t) < \infty$. Мы назовем ее <u>информационно-устойчивой</u>, если существует информационно-устойчивая последовательность пар $(\eta^t, \tilde{\eta}^t)$, связанных передающим устройством (Q^t, V^t) такая, что

$$\lim_{t \to \infty} \frac{J(\eta^t, \tilde{\eta}^t)}{C(Q^t, V^t)} = 1 \qquad (20)$$

Есть много оснований надеяться, что свойство последовательности передающих устройств быть при $C(Q^t, V^t) \to \infty$ информационно-устойчивой является столь же общим, как и свойство информационной устойчивости последовательностей случайных величин, однако уже полученные математические результаты являются здесь много менее широкими. Автором [21] доказана информационная устойчивость последовательности возрастающих по длине отрезков однородного канала без

VIII памяти. Используя теорию сумм независимых величин, повидимому нетрудно было бы получить общие (быть может необходимые и достаточные) условия информационной устойчивости для каналов без памяти в неоднородном случае, а также для "схемы серий".

Многочисленные исследования [99], [34], [35], [12], посвященные решению, поставленной Хинчиным задачи о совпадении эргодической и обычной пропускной способностей, можно интрепретировать как доказательство информационной устойчивости последовательности отрезков канала с конечной памятью в смысле Фейнстейна-Хинчина. Из результатов работы [8] можно, используя приводимую ниже (см. 2.10) теорему Ху-Го-Дина, вывести косвенным образом информационную устойчивость отрезков однородных каналов с конечной памятью и дискретными пространствами состояний при некоторых ограничениях на переходную функцию q_0. Кажется интересным дать для этого случая прямое доказательство информационной устойчивости и установить (в стиле эргодической теории однородных конечных цепей Маркова) условие, необходимое и достаточное для информационной устойчивости при любом начальном состоянии.

IX

Для недискретного случая единственным общим результатом является результат Пинскера, доказавшего: для того, чтобы последовательность гауссовских передающих устройств была информационно-устойчивой, необходимо и достаточно, чтобы $C(Q^t, V^t) \to \infty$. Здесь очень актуальным кажется вопрос о получении широких достаточных условий информационной устойчивости отрезков каналов с обобщениями на случай непрерывного времени и неоднородный по времени (быть может, почти-периодический).

X

2.8. Последовательность сообщений (p^t, W^t) ; с $H(p^t, W^t) < \infty$ назовем информационно-устойчивой, если существует информационно-устойчивая последовательность пар $(\xi^t, \tilde{\xi}^t)$;, образующих сообщение (p^t, W^t) ; такая, что

$$\lim_{t\to\infty} \frac{\mathcal{J}(\xi^t, \tilde{\xi}^t)}{H(p^t, W^t)} = 1 \qquad (21)$$

← Об условиях информационной устойчивости последовательности сообщений известно меньше, чем для передающего устройства.

В случае сообщений с условием полного вопроизведения вопрос сводится к информационной устойчивости пар величин $(\xi^t, \tilde{\xi}^t)$, где ξ^t имеет распределение p^t и может быть обычно решен на основе цитированных ранее результатов. В нашей работе [21] доказана информационная устойчивость последовательности однородных по времени сообщений с покомпонентным условием точности воспроизведения и независимыми компонентами при числе компонент $n \to \infty$. Здесь акту-

XI альна проблема, аналогичная проблеме VIII для передающих устройств. Пинскером доказано, что для информационной устойчивости последовательности гауссовских сообщений необходимо и достаточно, чтобы $H(p^t, W^t) \to \infty$. В дипломной работе студ. Герман (МГУ) доказана информационная устойчивость последовательности сообщений с покомпонентными условиями точности воспроизведения, когда процесс на входе является стационарным, эргодическим процессом, принимающим конечное чис-

XII ло значений. Здесь актуальны обобщения на случай непрерывного множества значений сообщения и непрерывного времени.

2.9. Введем следующее определение. Пусть $\mathcal{U} \subset R^n$, где R^n есть n-мерное евклидово пространство с точками $\bar{x} = (x_1, \dots, x_n)$. Нам будет удобно принять за расстояние между точками $\bar{x}' = (x_1', \dots, x_n')$, $\bar{x}'' = (x_1'', \dots, x_n'')$ число $r(\bar{x}', \bar{x}'') = \max_i |x_i' - x_i''|$. Обозначим через $[\mathcal{U}]_\varepsilon$, где $\varepsilon \geq 0$ совокупность точек $\bar{x} \in R^n$ таких, что для некоторых $\bar{\bar{x}} \in \mathcal{U}$ расстояние $r(\bar{x}, \bar{\bar{x}}) \leq \varepsilon$. Через (Q, V_ε) мы будем обозначать передающее устройство, в котором в определение (2) множество $\bar{V}$ заменено на $[\bar{V}]_\varepsilon$, а через (p, W_ε) будем обозначать сообщение, в котором множество $\bar{W}$ заменено на $[\bar{W}]_\varepsilon$.

2.10. Мы можем теперь сформулировать следующую теорему Шеннона: Пусть заданы информационно-устойчивые последовательности передающих устройства (Q^t, V^t) и сообщений (p^t, W^t) такие, что функции $\pi_i^t(\cdot,\cdot)$ и $\rho_j^t(\cdot,\cdot)$, $i=1,\dots,N^t$, $j=1,\dots,M^t$ ограничены равномерно по i и t и что при любом $a>0$ $N^t = o(e^{aC(Q^t,V^t)})$, $M^t = o(e^{aH(p^t,W^t)})$. Пусть $H(p^t,W^t)\to\infty$ и нижний предел

$$\varliminf_{t\to\infty} \frac{C(Q^t,V^t)}{H(p^t,W^t)} > 1. \qquad (22)$$

Тогда для любого $\varepsilon>0$ существует столь большое t_0, что при всех $t\geqslant t_0$ сообщение (p^t, W_ε^t) может быть передано через передающее устройство (Q^t, V_ε^t).

Мы будем называть эту теорему прямой теоремой Шеннона.

Обсудим этот результат. Прежде всего заметим, что не всегда можно заменить в формулировке теоремы W_ε^t на W^t и V_ε^t на V^t. Это ясно видно, например, для сообщений с условием полного воспроизведения, так как возникающая хотя бы с малой вероятностью ошибка при передаче по любому каналу с шумами приведет здесь к неполному совпадению сообщения на входе и выходе. ~~Хотя~~ Зато обратная теорема Шеннона о невозможности передачи показывает, что сообщение (p^t, W_ε^t) может быть передано через передающее устройство (Q^t, V_ε^t) лишь если $\frac{C(Q^t,V_\varepsilon^t)}{H(p^t,W_\varepsilon^t)} \geqslant 1$. Так как обычно это отношение непрерывно зависит от ε, то из обратной теоремы Шеннона следует, что в широком классе случаев условие (22) (со знаком $>$, замененным на $\geqslant$) также и необходимо для выполнения утверждения сформулированной теоремы. Именно в такой форме обратная теорема Шеннона дается в большинстве предыдущих работ. В [21] указывается условия, при которых в сформулированной теореме можно все же заменить W_ε^t на W^t и V_ε^t на V^t.

Ограничение на числа функций $\pi_i^t(\cdot,\cdot)$ и $\rho_j^t(\cdot,\cdot)$ является достаточно слабым, и выполняется во [illegible] интересных частных случаях. Наоборот, требование ограниченности функций $\pi_i^t(\cdot,\cdot)$ и $\rho_j^t(\cdot,\cdot)$ является крайне жестким. В [21] подробно исследуются дополнительные условия, которые нужны для того, чтобы заменить требования ограниченности. Они тесно связаны с требованиями информационной-устойчиво- сти и напоминают условия Ляпунова для центральной пре- дельной теоремы. Кажется интересным исследовать их выпол-

XIII нение в конкретных ситуациях / см. 2.7, 2.8 /. В [21] этот вопрос рассмотрен частично для случаев канала без памяти и сообщения с покомпонентным условием точности и незави- симыми компонентами.

Основными являются условия информационной-устойчиво- сти последовательностей передающих устройств и сообще- ний. Недавно Ху Го Дином было доказано, что требование, очень близкое к утверждению об информационной устойчиво- сти последовательности передающих устройств необходимо для того, чтобы при достаточно больших t через пере- дающее устройство могло быть передано любое информацион- но-устойчивое сообщение с условием полного воспроизведе- ния, для которого верно (22). Результат Ху Го Дина пока- зывает, что введение условия информационной устойчивости передающих устройств является неизбежным. Кажется интерес-

XIV ным попытаться получить аналогичный результат и для тре- бования информационной устойчивости последовательности сообщений.

Приведенная теорема Шеннона включает в себя как частный случай / если не учитывать менее существенные вариации в формулировках/ все ранее опубликованные теоремы такого типа / см. [14],[36],[67],[64] /. Однако в перечисленных работах, кроме самой теоремы Шеннона, исследуется вопрос об условиях информационной устойчивости рассматриваемых в них передающих устройств и сообщений, выделяемый при нашем плане изложения в отдельную проблему.

2.11. В случае, когда последовательность передающих устройств - это последовательность отрезков однородного канала, а последовательность сообщений - последовательность однородных сообщений с растущим числом компонент, предпочитают часто несколько иную формулировку основных теорем Шеннона.

Введем в связи с этим следующие понятия. Пусть $(Q^n_{p_0}, V^n)$ - передающее устройство, задаваемое отрезком длины n однородного канала с начальным распределением p_0. Назовем средней пропускной способностью канала

$$\bar{C} = \lim_{n\to\infty} \frac{1}{n} C(Q^n_{p_0}, V^n), \qquad (23)$$

если этот предел существует для любого начального распределения p_0 и не зависит от этого распределения. Пусть далее $\{p^n, W^n\}$ - однородное по времени сообщение с n компонентами. Мы будем называть скоростью создания сообщения предел

$$\bar{H} = \lim_{n\to\infty} \frac{1}{n} H(p^n, W^n) \qquad /24/$$

эти определения легко обобщаются на случай непрерывного времени. Из сформулированных выше теорем Шеннона немедленно следует, что если $\bar{H} > \bar{C}$, то при достаточно больших n сообщение $\{p^n, W^n\}$ не может быть передано че- [redacted] рез передающее устройство $(Q^n_{P_0}, V^n)$). Если, наоборот, $\overline{C} > \overline{H}$ и выполнены требования информационной устойчивости, то при достаточно больших n сообщение (P^n_0, W^n_0)) может быть передано через передающее устройство $(Q^n_{P_0}, V^n)$).

Иногда предпочитают несколько иной вариант этих определений. Пусть $\mathcal{Z} = \{\zeta_n, \tilde{\zeta}_n\}$ совместно-стационарные процессы. Скоростью передачи информации назовем

$$\overline{J}(\mathcal{Z}) = \lim_{n\to\infty} \frac{1}{n} J((\zeta_1,\dots,\zeta_n),(\tilde{\zeta}_1,\dots,\tilde{\zeta}_n)). \qquad (25)$$

(Условия для существования предела (25), а также другие варианты этого определения исследованы Пинскером). Тогда естественно предположить, что

$$\overline{C} = \sup J(\mathcal{H}) \qquad (26)$$

где верхняя грань берется по всем процессам $\mathcal{H} = \{\eta_n, \tilde{\eta}_n\}$ таким, что при любом n величины $(\eta_1,\dots,\eta_n)$ и $(\tilde{\eta}_1,\dots,\tilde{\eta}_n)$ связаны отрезком канала длины n при некотором начальном распределении для состояния канала, и что

$$\overline{H} = \inf J(\Xi) \qquad (27)$$

где нижняя грань берется по всем процессам $\Xi = \{\xi_n, \tilde{\xi}_n\}$ таким, что при любом n величины $(\xi_1,\dots,\xi_n)$ и $(\tilde{\xi}_1,\dots,\tilde{\xi}_n)$ образуют сообщение с n компонентами. Вывод равенства (26) содержится для каналов с конечной памятью по Хинчину-Файнстейну в [99] и [34]. Проблема установления общих условий
XIV для его выполнения, а также еще почти совсем неизученная проблема установления условий для равенства (27) повидимому тесно связаны с проблемами IX, X, XII получения условий информационной устойчивости соответствующих канала и сообщения.

2.12. Приведенные формулировки теоремы Шеннона отражают математически реальную картину передачи информации, длящейся некоторый большой, но конечный отрезок времени. Но в большинстве предыдущих работ вводилась иная математическая идеализация, при которой передача информации мыслится длящейся бесконечно долго. Мы не будем приводить соответствующие определения, поскольку для простейших случаев они даны в хорошо известных работах, а в общем случае громоздки и содержатся в [21], разд. 1.8).

Для простейших случаев, уже изученных в литературе ранее, переход от теоремы Шеннона в приведенной выше формулировке к соответствующей теореме для бесконечной передачи почти тривиален, однако в общем случае это уже не так. Опишем наглядно пример, раз"ясняющий возникающие при этом трудности. Пусть подлежащее передаче сообщение $\xi^t \equiv \xi$, где $-\infty < t < \infty$, а ξ - случайная величина с непрерывным распределением, и условие точности воспроизведения является условием полного воспроизведения. Скорость создания такого сообщения бесконечна. Однако существуют методы кодирования и декодирования, позволяющие передать его через передающее устройство с сколь угодно малой ненулевой средней пропускной способностью. Для этого, взяв монотонную последовательность моментов времени $T_1 > T_2 > \ldots$, $T_n \to -\infty$, за отрезок $[T_{n+1}, T_n]$ передадим сведения о ξ_t с точностью до 2^{-n}. Выбрав $T_{n+1} - T_n$ достаточно большим, это можно сделать со сколь угодно малой вероятностью ошибки. К любому моменту времени t уже до этого момента на выходе будут получены сведения о ξ_t со сколь угодно большой точностью и малой вероятностью ошибки. Таким образом, здесь $\overline{H} > \overline{C}$, а передача за бесконечное время тем не менее возможна. Можно привести также (существенно более сложный) пример аналогичной ситуации, где в котором

XV Кажется интересным разобраться до конца в этом круге вопросов. Для этого, быть может, понадобится несколько реформировать определение энтропии и пропускной способности (ср. различные определения скорости передачи информации в [53]).

3. Вычисление пропускной способности, W -энтропии и оптимальных распределений.

3.1. Поскольку решение проблемы Шеннона дается через пропускную способность передающего устройства и W -энтропию сообщения, важно научиться вычислять явно эти величины в конкретных случаях. Кроме того интересна тесно связанной с предыдущим вопрос о вычислении оптимальных распределений на входе передающего устройства (т.е. распределений P в пространстве сигналов на входе таких, что если η имеет распределение p и пара $(\eta, \tilde{\eta})$ связаны передающим устройством (Q,V)), то

$$\mathcal{J}(\eta,\tilde{\eta}) = C(Q,V) \tag{28}$$

и оптимальных распределений для сообщения (т.е. совместных распределений $r(\cdot,\cdot)$) в произведении пространств значений сообщения на входе и на выходе таких, что пары $(\xi,\tilde{\xi})$) с распределением $r(\cdot,\cdot)$ образуют сообщение (p,W) и

$$\mathcal{J}(\xi,\tilde{\xi}) = H(p,W). \tag{29}$$

Важность оптимальных распределений об"ясняется тем, что из неравенства (15) видно, что если величины $\xi, \eta, \tilde{\eta}, \tilde{\xi}$

задают способ передачи сообщения (p,W) по передающему устройству (Q,V), то

$$\mathcal{J}(\xi,\tilde{\xi}) \le C(Q,V), \quad \mathcal{J}(\eta,\tilde{\eta}) \ge H(p,W) \qquad (30)$$

Отсюда следует, что когда энтропия сообщения близка к пропускной способности передающего устройства (т.е. эта пропускная способность используется почти полностью), методы кодирования и декодирования надо выбирать так, чтобы разности $\mathcal{J}(\xi,\tilde{\xi}) - H(p,W)$ и $C(Q,V) - \mathcal{J}(\eta,\tilde{\eta})$ были малы. Отсюда, естественно, сделать вывод, что распределение сигнала на входе η и значений сообщения $\xi, \tilde{\xi}$ близки к оптимальным, и значит, значение оптимальных распределений подсказывает как должен быть выбран метод передачи. С точки
XVI зрения математика утверждение о том, что из малроти разностей $\mathcal{J}(\xi,\tilde{\xi}) - H(p,W)$ и $C(Q,V) - \mathcal{J}(\eta,\tilde{\eta})$ следует близости соответствующих распределений, далеко не очевидна, поскольку пространство распределений не компактно, а информация — не непрерывный функционал. Кажется тем не менее, что это верно в достаточно широких предположениях.

3.2. В соответствии с определением пропускной способности (16) задача ее вычисления является задачей о вычислении максимума функционала (а в случае дискретного пространства Y -функции нескольких переменных). Ее специфика связана с особым аналитическим видом рассматриваемого функционала, позволяющим упростить решение, а также с тем, что на область изменений его аргумента положены ограничения типа неравенств (ограничение V и требование неотрицательности распределений вероятностей). Общая теория вариационных задач с такими ограничениями еще недостаточно разработана (см. [5]).

Получить явное выражение для пропускной способности удается лишь в исключительных случаях, однако в работах Шеннона и особенно Мууроги ([80], [61], [62]). разработан алгоритм, сводящий задачу к решению трансцендентного уравнения. Эти исследования проведены на уровне "физической строгости"
XVII и кажется важным (в особенности для более сложного недискретного случая) завершить их, разобрав подробно вопросы существования, исследовав особые случаи и т.п. В работе Шеннона [82] дана для дискретного случая геометрическая трактовка задачи и, в частности, показано, что множество оптимальных распределений выпукло. Интересно создать методы вычисления крайних точек этого множества, а также перенести исследование на общий случай. XVIII

XIX Аналогичные проблемы встают в связи с вычислением энтропии сообщений и оптимальных распределений для сообщения. Однако здесь не сделано еще почти ничего (если не считать беглых замечаний в [85]).

3.3. Для случая, когда передающее устройство является отрезком канала без памяти нетрудно показать (см. [21] или [33]), что среди оптимальных распределений на входе имеется распределение с независимыми компонентами. Отсюда легко выводится, что если C_n -пропускная способность отрезка длины n , то

$$C_n = n C_1 \tag{31}$$

Однако вычисление величины C_1 является обычно так же трудной задачей. Для гауссовского канала без памяти с аддитивным шумом верна формула Шеннона

$$C_n = n \log\left(1 + \frac{P_s}{P_n}\right), \tag{32}$$

а оптимальное распределение является гауссовским.

Аналогично, для сообщения с покомпонентным условием точности и независимыми компонентами среди оптимальных распределений имеются (см. [21]) распределения, для которых пары $(\xi_i, \tilde{\xi}_i)$ независимы. Отсюда снова следует, что если H_n - энтропия сообщения с n компонентами, то

$$H_n = n H_1 \tag{33}$$

Особенно простой является следующая формула для однородного гауссовского сообщения с условием ограниченности средне-квадратичной ошибки

$$H_n = n \log \frac{P_d}{P_c} \tag{34}$$

где P_d - средне-квадратичная ошибка, а P_c - дисперсия компоненты сообщения на входе.

3.4. В более общих ситуациях явных формул для пропускных способностей и W - энтропий обычно не удается и в связи с этим ставится несколько более простая задача вычисления средней пропускной способности $\overline{C}$ и скорости создания сообщения $\overline{H}$ (ср. разд. 2.11).

Проблема вычисления величин $\overline{C}$ и $\overline{H}$ уже нашла достаточно полное решение для важнейших с прикладной точки зрения гауссовских передающих устройств и сообщений. Дело в том, что как нетрудно показать (см. Пинскер [71]) в этих случаях среди оптимальных распределений всегда есть гауссовские распределения. А для гауссовских распределений можно дать (см. [80], [40]) простое выражение для информации в конечномерном случае, допускающее обобщения и на бесконечномерный случай (в частности в [40] развит метод вычисления информации для некоторых классов гауссовских процессов). Общие формулы для скорости передачи информации, средней пропускной способности, скорости создания сообщения через соответствующие спектральные плотности получены Пинскером [70], [71]. Частично его результаты повторены в [74]. Методы решения подобных задач тесно связаны с аналитическими методами корреляционной теории стационарных процессов.

Наиболее популярной в технических приложениях является следующая формула Шеннона, органически вошедшая в мышление радиоинженеров: Если C_T - пропускная способность передающего устройства с конечной полосой пропускания и белым гауссовским аддитивным шумом для передачи за время T, то при $T \to \infty$

$$C_T \sim 2WT \log\left(1+\frac{P_s}{P_n}\right). \qquad (35)$$

Для других классов процессов дело обстоит хуже. Так, например, для часто используемых в радиотехнике релеевских процессов (эти процессы возникают при прохождении узкополосных гауссовских процессов через детекторы для выделения огибающей (см. напр. [16]) проблема приводит сразу к серьезным трудностям (кое-что по этому поводу сделано в [62]).

XX

3.5. Легко выводится простая формула скорости создания сообщения с условием полного воспроизведения, если процесс на входе является однородной конечной цепью Маркова (см. [80], [13]). Однако, если только заменить здесь цепь Маркова на процесс, являющийся функцией от цепи Маркова, задача становится неизмеримо сложнее. Ее остроумное решение предложено Блеквеллом [6]. Наиболее актуальным кажется здесь вопрос

XXI о том, нельзя ли обобщить это решение на случай, когда условие точности является произвольным покомпонентным или аддитивным условием. Возможно, что аналогичные методы можно применить

XXII к еще совсем неисследованной задаче о нахождении средней пропускной способности канала с конечной памятью. В этом случае среди оптимальных процессов на входе (т.е. процессов $\mathcal{H}$, для которых достигается верхняя грань в [26], будут повидимому (см. [8]) процессы, являющиеся функциями от марковских процессов. Кажется интересным дать метод для вычис-

XXIII ления параметров таких процессов.

3.6. Трудность нахождения явных выражений для W-энтропий и пропускных способностей естественно приводит к мысли использовать малость некоторых параметров канала и сообщения.

Для случая, когда передающее устройство таково, что сигнал на выходе мало отличается от сигнала на входе и для случая, когда сообщение таково, что значение его на выходе близко к его значению на входе, задача оказывается тесно связанной с невероятностной теорией ε-энтропий и ε-емкостей /см. [56], доб.2). Другим полезным ограничением может служить предположение о том, что канал является каналом со случайным параметром. Естественно предположить, что в этом случае средняя пропускная способность

$$\bar{C} = \sup_{\mathcal{Z}} \int_{B} \bar{J}(\mathcal{Z}_{\beta})\, \bar{P}(d\beta) \tag{36}$$

где $\bar{P}(.)$ - распределение параметра $\beta(\varphi)$, а $\bar{J}(\mathcal{Z}_{\beta})$ -

скорость передачи информации для пары процессов $(\xi_n, \tilde{\xi}_n^{\beta})$ таких, что $(\xi_1,\dots,\xi_n)$ и $(\tilde{\xi}_1^{\beta},\dots,\tilde{\xi}_n^{\beta})$ связаны соответствующим условным каналом с параметром β, а верхняя грань берется по всем процессам на входе $\{\xi_n\}$, для которых существует процесс на входе $\{\tilde{\xi}_n\}$ такой, что пары $(\xi_1,\dots,\xi_n)$, $(\tilde{\xi}_1,\dots,\tilde{\xi}_n)$. связаны исследуемым каналом. Из работ Цыбакова ([101], [102], [103]) следует, что формула (36) позволяет вычислять пропускную способность широкого класса физически-реальных каналов распространения радиоволн, в частности при наличии флюктуаций фазы. В работе автора [18] формула (36) доказана для одного частного случая, причем доказательство легко распространяется на общий случай, лишь бы пространство B значений параметра было конечным. Поскольку однако в упомянутых приложениях формулы (36) приходится иметь дело с величиной β, имеющей непрерывное распределение, интересно вывести

XXIV формулу (36) и для этого случая.

Отметим однако, что канал со случайным параметром является неэргодическим при любой разумной трактовке этого термина и поэтому для него неверна прямая теорема Шеннона. Это не лишает однако физического смысла значение $\overline{C}$, поскольку его можно трактовать как приближенное значение для пропускной способности канала с медленно меняющимся параметром, к которым прямая теорема Шеннона применима.

3.7. Из важных, но не рассматривавшихся здесь направлений исследования, отметим работы по вычислению пропускной способности многолучевых каналов [65] и по статистической оценке теоретико-информационных параметров реальных каналов и сообщений ([4], [19], [57]).

4. Исследование оптимальных кодов. 4.1. Естественным образом встает вопрос как наиболее простым образом построить методы кодирования и декодирования, существование которых гарантирует теорема Шеннона. Мы рассмотрим эту проблему лишь для наиболее простого случая, когда сообщение является сообщением с условием полного воспроизведения таким, что пространство X значений сообщения на входе состоит из S элементов $E_1, \ldots, E_S$ и $P_\xi(E_i) = \frac{1}{S}$, $i = 1, \ldots, S$. Это сообщение мы будем обозначать $(P, \mathfrak{M}^S)$. Такой случай особо важен, поскольку ход доказательства теоремы Шеннона для произвольного сообщения (см. [21]) подсказывает, что решение проблемы для этого частного случая войдет как непременная составляющая в ее общее решение. Мы не будем останавливаться здесь на интересной серии исследований (см. например [42]), посвященной так называемым неравномерным кодам, которые можно интрепретировать как исследования эффективных методов кодирования для иного класса сообщений.

4.2. Пусть задано передающее устройство (Q, V). Через $e(Q, V, S)$ будем обозначать наименьшее ε такое, что сообщение $(P, \mathfrak{M}^S_\varepsilon)$ может быть передано через передающее устройство (Q, V). Поскольку последовательность сообщений $\mathfrak{M}^S$ является при $S \to \infty$ информационно-устойчивой с энтропией $\log S$, то если

$$\lim_{t \to \infty} \frac{\log S^t}{C(Q^t, V^t)} < 1 \tag{37}$$

и для последовательности $C(Q^t, V^t)$ удовлетворяются предположения прямой теоремы Шеннона, выполнено соотношение

$$\lim_{t \to \infty} e(Q^t, V^t, S^t) = 0. \tag{38}$$

Из обратной теоремы Шеннона вытекает, что если наоборот

$$\lim_{t\to\infty} \frac{\log S^t}{C(Q^t,V^t)} > 1, \tag{39}$$

то

$$\lim_{t\to\infty} e(Q^t,V^t,S^t) > 0. \tag{40}$$

Используя построения, применяемые при доказательстве прямой теоремы Шеннона, нетрудно показать, что если последовательность (Q^t,V^t) информационно-устойчива, то (40) можно заменить на более сильное равенство (41)

$$\lim_{t\to\infty} e(Q^t,V^t,S^t) = 1 \tag{41}$$

Различие между формулировками (40) и (41) подчеркивалось Вольфовицем ([97]).

4.3. Если равенство (41) достаточно исчерпывающим образом описывает асимптотику $e(\cdot,\cdot,\cdot)$ при условии (39), то наоборот, в случае выполнения (37) хочется уточнить соотношение (38), оценив скорость стремления $e(\cdot,\cdot,\cdot)$ к нулю. Такая оценка особенно важна, если (Q^n,V^n) - отрезок длины n однородного канала, поскольку если $e(Q^{n_0},V^{n_0},S^{n_0})$ достаточно мало, то кодирование для передачи по длительно действующему каналу можно проводить по отдельным блокам длины n_0 (ср. аналогичные построения в [33], [14]) и значит, чем меньше n_0, тем проще будет подобный метод кодирования. Повидимому, при очень широких предположениях из (37) следует, что при некотором $a>0$

$$e(Q^t,V^t,S^t) = o\left(2^{-aC(Q^t,V^t)}\right). \tag{42}$$

← Во всяком случае, прослеживая доказательство прямой теоремы Шеннона (см. [21]), легко заметить, что это будет так, если предположение информационной устойчивости последовательности (Q^t, V^t) заменить на более сильное предположение о том, что существует последовательность пар $(\xi^t, \tilde{\xi}^t)$, связанных передающим устройством (Q^t, V^t), такая, что при любом $\varepsilon > 0$ и некотором $d(\varepsilon) > 0$

$$\lim_{t\to\infty} \mathcal{P}\left\{ \left| \frac{i_{\xi^t \tilde{\xi}^t}(\xi^t, \tilde{\xi}^t)}{C(Q^t, V^t)} - 1 \right| > \varepsilon \right\} = o\left(2^{-d(\varepsilon) C(Q^t, V^t)}\right) \quad (43)$$

Для канала без памяти условие (43) следует, если предположить, что оптимальное распределение на входе является "достаточно хорошим" из известных теорем о больших уклонениях сумм независимых величин (см. [15]), (таким путем Файнстейн [32] доказал утверждение (42) для однородных конечных каналов без памяти. Для каналов с конечным числом состояний, утверждение (43) может быть, повидимому, выведено из известных оценок больших уклонений для сумм величин, связанных в цепь Маркова, однако в более общей ситуации нужны отдельные исследования.
XXV

4.4. Предположим теперь, что (Q^n, V^n) - отрезок длины n однородного канала. Положим далее $S^n = [2^{nH}]$. Условие (37) означает теперь, что

$$H < \bar{C} \quad (44)$$

где $\bar{C}$ - средняя пропускная способность канала. Исследования, проведенные в простейших случаях, позволяют надеяться, что при $n \to \infty$ и некоторых $a, b > 0$

$$e(Q^n, V^n, [2^{nH}]) \asymp n^a 2^{-bn} \qquad (45)$$

(Здесь обозначение $x_n \asymp y_n$ подразумевает, что $0 < \underline{\lim}_{n\to\infty} \frac{x_n}{y_n} \le \overline{\lim}_{n\to\infty} \frac{x_n}{y_n} < \infty$). Более точную асимптотику вида

$$e(Q^n, V^n, [2^{nH}]) \sim c\, n^a 2^{-bn} \qquad (46)$$

где c -константа, получить видимо невозможно (хотя это и не доказано пока что ни для одного случая). Дело в том, что при исследовании вероятности $e(\cdot,\cdot,\cdot)$ с такой высокой точностью существенными оказываются арифметические свойства числа n .

XXVI Кажется интересным доказать существование констант a и b , но еще важнее, конечно, научиться их вычислять. Эта задача оказывается трудной даже для простейших конечных каналов без памяти. Автор рассмотрел эту задачу для каналов без памяти таких, что матрица вероятностей перехода $Q_0=(q_{ij})$ обладает следующим свойством симметричности: каждая ее строка является перестановкой любой другой строки и каждый столбец является перестановкой любого другого столбца. Оказалось, что для таких каналов

$$b = \log R(\bar{k}) + (1-\bar{k})\log m(\bar{k}) + \log N$$
$$a = -\frac{1}{2\bar{k}}, \qquad (47)$$

где

$$R(k) = \frac{N}{M}\sum_{i=1}^{M}(q_{i1})^k, \quad m(k) = \frac{d\log R(k)}{dk}$$
$$\log R(\bar{k}) - \bar{k}\, m(\bar{k}) = -H, \qquad (48)$$

M – число элементов в пространстве сигналов на входе, а $\mathcal{N}$ – в пространстве сигналов на выходе, если

$$H \geqslant H_{crit} = \frac{1}{2} m\left(\frac{1}{2}\right) - \log R\left(\frac{1}{2}\right).$$

Частным случаем рассмотренных каналов является симметричный бинарный канал с матрицей $Q = \begin{pmatrix} q_1 & q_2 \\ q_2 & q_1 \end{pmatrix}$, $q_1 + q_2 = 1$. Для этого канала Элайас [26], [27] изучил интересующую нас задачу x/. В случае $H < H_{crit}$ удается получить лишь несовпадающие верхнюю и нижнюю оценку для b. Задача точного

XXVII вычисления b оказывается здесь даже для простейшего симметричного бинарного канала очень трудной и ее решение, видимо, возможно лишь с привлечением каких-либо новых идей. Элайас [26] исследовал также случай бинарного канала со стиранием, т.е. канала с матрицей $\begin{pmatrix} q_1 & 0 & q_2 \\ 0 & q_1 & q_2 \end{pmatrix}$, $q_1 + q_2 = 1$ не являющегося симметричным в нашем смысле. Кажется, что

XXVIII этот результат Элайса может быть обобщен на некоторый класс небинарных каналов. Однако уже для несимметричного бинарно-

XXIX го канала задача нахождения параметров a и b становится снова очень трудной и подходы к ней пока что не видны (грубо говоря там H_{crit} совпадает со средней пропускной способностью $\bar{C}$, так что трудности здесь те же, что и для симметрического бинарного канала при $H < H_{crit}$). В общем случае конечного канала без памяти удается получить лишь некоторые оценки (см. [83]).

Шеннон [86] исследовал задачу для важного класса гауссовских каналов без памяти с аддитивным шумом, где она снова нашла полное решение лишь для $H \geqslant H_{crit}$. Неясно, нас-

XXX колько его результаты допускают обобщения на иные гауссовские каналы. Повидимому результаты, полученные для симметрич-

x) Однако его выражение для a ошибочно.

ных конечных каналов без памяти могут быть сообщены на некоторые каналы с непрерывным множеством сигналов, образующими окружности $|z|=1$ и такими, что их переходные вероятности инвариантны при поворотах (такие каналы могут хорошо описывать передачи при помощи фазовой модуляции).

XXXI

4.5. Во всех цитированных работах для построения методов кодирования с наименьшей ошибкой используется метод случайного кодирования. Состоит он в следующем. Будем рассматривать лишь нерандомизированное кодирование. В соответствии со сделанными выше замечаниями (см. разд. 1.10) - это не есть существенное ограничение общности. Кодирование можно тогда задавать как функцию $f(E_i)$. Кодом $\mathfrak{K}(s)$ мы будем называть набор значений $f(E_1), \ldots, f(E_s)$. Через $e(Q,V,\mathfrak{K}(s))$ мы будем обозначать наименьшее ε такое, что сообщение $(p, \mathfrak{M}_\varepsilon^s)$ может быть передано через передающее устройство (Q,V) с использованием кодирования $\mathfrak{K}(s)$ и некоторого метода декодирования. Предположим теперь, что нам задано некоторое распределение вероятностей $r(.)$ на пространстве сигналов на входе (X, S_X) и система независимых случайных величин $\varphi_1, \ldots, \varphi_s$ - с распределением $r(.)$. Мы будем называть $\mathfrak{K}(s)=(\varphi_1,\ldots,\varphi_s)$ случайным кодом, а математическое ожидание

$$\bar{e}(Q,V,s) = E\{e(Q,V,\mathfrak{K}(s))\} \tag{49}$$

средней вероятностью ошибки для распределения $r(.)$. Очевидно, что $\bar{e}(Q,V,s) \geq e(Q,V,s)$. Все известные и перечисленные выше асимптотические оценки сверху для вероятности $e(Q,V,s)$ основаны на этом неравенстве, однако исследование средней вероятности ошибки имеет и самостоятельный интерес, поскольку она характеризует вероятность ошибки для типичного кода (Есть основания надеяться, что в широком классе случаев верен следующий закон больших чисел: при $n \to \infty$

XXXII
$$\frac{e(Q^n, V^n, \mathcal{D}([2^{nH}]))}{\bar{e}(Q^n, V^n, [2^{nH}])} \to 1 \quad \text{(по вероятности)} \tag{50}$$

Кажется также интересным исследовать подробнее асимптотическое распределение для $e(Q^n, V^n, \mathcal{D}([2^{nH}]))$ при $n \to \infty$ В качестве $\tau(.)$ берут чаще всего оптимальное распределение сигнала на входе. При этом из обычного доказательства теоремы Шеннона (см. [21]) видно, что из условия (43) вытекает не только соотношение (42), но и более сильное утверждение

$$\bar{e}(Q^t, V^t, S^t) = o\left(2^{-aC(Q^t, V^t)}\right). \tag{51}$$

XXXIII
Повидимому, в широком классе случаев

$$\bar{e}(Q^n, V^n, [2^{nH}]) \asymp n^{\bar{a}} 2^{-\bar{b}n}. \tag{52}$$

Вычислены значения $\bar{a}$ и $\bar{b}$ пока лишь для тех же случаев, когда (см. выше) исследовались константы a и b . Оказалось, что при $H \geqslant H_{crit}$ $a = \bar{a}$, $b = \bar{b}$. Это означает, что при достаточно высоких скоростях передачи информации случайный код асимптотически (с точностью до констант) так же хорош, как и оптимальный, что позволяет (см. [37]) строить "хорошие" коды методом Монте-Карло.

Отметим, однако, что для гауссовского канала Шеннону [86] пришлось брать в качестве $\tau(.)$ не оптимальное распределение, а некоторое лишь асимптотически оптимальное распределение (вместо распределения n независимых нормальных величин со средним 0 и дисперсией P_s, ему пришлось взять равномерное распределение на n -мерной сфере радиуса $\sqrt{nP_s}$). Интересно, не может ли подобный прием привести

XXXIV к уменьшению $\bar{a}$ и $\bar{b}$ и в других случаях. При $H < H_{crit}$ явные выражения для $\bar{a}$ и $\bar{b}$ также найти удается, например, для описанных выше симметрических каналов

$$\bar{a} = -\frac{1}{2}, \ \bar{b} = H + 2\log_2 k\left(\tfrac{1}{2}\right) + \log_2 N. \qquad (53)$$

Интересно, что во всех исследованных случаях H_{crit} оказалось точкой излома производной функции $\bar{b}(H)$. Повидимому при $H < H_{crit}$ $b < \bar{b}$, однако это не доказано даже

XXXV для симметричного бинарного канала (ср. пробл. XXYII). Задача вычисления $\bar{a}$ и $\bar{b}$ проще, чем задача вычисления констант a и b и кажется реальным найти $\bar{a}$ и $\bar{b}$ в достаточ-

XXXVI но широком классе случаев.

4.6. Трудность практического осуществления передач и теоретического исследования произвольных кодов привели к выделению особого класса групповых кодов (см. [44], [89], [10], [91]).

Предположим, что рассматривается конечный канал без памяти такой, что число элементов в пространстве сигналов на входе $M = q^k$, где q - простое число, а k - целое число. Отождествим Y_0 с прямым произведением k циклических групп порядка q, а Y^n будем рассматривать как группу, являющуюся прямым произведением n групп Y_0.

Код $\mathfrak{A}(S)$ назовем <u>групповым</u>, если он образует подгруппу группы Y^n. Обозначим через $\tilde{e}(Q,V,S)$ наименьшее значение вероятности $e(Q,V,\mathfrak{A}(S))$, когда минимум берется по всем групповым кодам. Оказывается, что для широкого класса однородных каналов без памяти, обладающих свойством симметричности переходной матрицы, а также для симметричного бинарного канала и бинарного канала со стиранием (см. [27]) при $n\to\infty$ и $H \geq H_{crit}$

$$e(Q^n,V^n,[2^{nH}]) \sim \tilde{e}(Q^n,V^n,[2^{nH}]), \tag{54}$$

и таким образом наилучший групповой код асимптотически столь же хорош, как и наилучший их всех кодов. Этот факт является следствием того, что при некотором естественном определении понятия случайного группового кода среднее значение вероятности ошибки по всем групповым кодам асимптотически совпадает при всех H с $\tilde{e}(Q^n,V^n,[2^{nH}])$. Таким образом, почти все групповые коды при $H \geq H_{crit}$ устроены как оптимальный код (чтобы уточнить это утверждение, хотелось бы и для групповых кодов получить равенство
XXXVII типа (50)). При $H < H_{crit}$ асимптотическое поведение
XXXVIII $\tilde{e}(Q^n,V^n,[2^{nH}])$ неизвестно даже для бинарного симметрического канала. Результаты, описанные выше, можно повидимому распространить на несколько более широкий класс ка-
XXXIX налов без памяти (ср. пробл. XXVIII). Очень интересным кажется вопрос о том, будут ли групповые коды асимптотически
XL оптимальными, хотя бы для простейших классов каналов с памятью.

Создание оптимальных алгебраических методов кодирования для $M \neq q^k$, повидимому, невозможно. Что касается каналов с непрерывным множеством состояний, то здесь, повидимому, возможны обобщения на случай, упомянутый в связи с пробл. XXXII, когда $\mathcal{Y}$ можно отождествить с группой по умножению комплексных чисел $|z|=1$. Пути к обобщению алгебраических методов кодирования, при котором они стали бы применимы, скажем, к гауссовским каналам без памяти в настоящее время также не видно.

4.7. Задача построения оптимальных кодов, т.е. задача создания относительно простых алгоритмов для задания кода $\mathfrak{K}(S)$ такого, что $e(Q^n, V^n, \mathfrak{K}(S))$ совпадает (или по крайней мере) близка к $e(Q^n, V^n, S)$ очень трудна. Ее удается решить лишь для отдельных небольших значений n и S. Она становится лишь немного легче, если ограничиваться рассмотрением групповых кодов (см. [89]). Ввиду этого введены и плодотворно изучаются иные невероятностные методы для оценки качества кода. Предположим, что (Q^n, V^n) — конечный канал без памяти. Тогда $\mathcal{Y}^n$ - это пространство последовательностей $y=(y_1,\dots,y_n)$, где y_i принимает значения $1,\dots,M$. Положив $\rho((y_1,\dots,y_n),(y_1',\dots,y_n'))$ равным числу индексов $i=1,\dots,n$ таких, что $y_i \neq y_i'$ мы превратим $\mathcal{Y}^n$ в метрическое пространство.

Z (абзац)

Мы назовем кодовым расстоянием кода $\mathfrak{K}=(f(E_1),\dots,f(E_S))$

$$d(\mathfrak{K}) = \min_{i\neq j} \rho(f(E_i), f(E_j)) \tag{55}$$

Интуиция подсказывает, что, вообще говоря, для каналов с симметрическими матрицами для кодов с большим кодовым расстоянием вероятность ошибки меньше. Нетрудно, однако, привести даже для бинарного канала примеры кодов таких, что $d(\mathfrak{K}) < d(\mathfrak{K}')$, но $e(Q^n, V^n, \mathfrak{K}) > e(Q^n, V^n, \mathfrak{K}')$.

Совпадение кодов оптимальных в смысле вероятности ошибки и в смысле кодового расстояния удается доказать лишь в весьма узком (см. [88]) классе случаев, когда существуют <u>плотно-упакованные коды</u>, т.е. когда все пространство Y^n можно представить как сумму непересекающихся сфер равного радиуса (плотно-упакованный код образуют центры этих сфер). Не доказано даже асимптотическое совпадение этих двух понятий оптимальности, хотя можно выдвинуть гипотезу, что если положить

$$d(n,s) = \max d(\mathfrak{K}(s)) \tag{56}$$

по всем кодам $\mathfrak{K}(s)$ об"ема S для передачи длины n, то из того, что для последовательности кодов $\mathfrak{K}([2^{nH}])$, $n = 1, 2, \ldots$

$$\frac{d(\mathfrak{K}([2^{nH}])}{d(n,[2^{nH}])} \to 1 \quad (n \to \infty) \tag{57}$$

вытекает, что

XLI

$$\frac{e(Q^n, V^n, \mathfrak{K}([2^{nH}])}{e(Q^n, V^n, [2^{nH}])} \to 1 \quad (n \to \infty) \tag{58}$$

Обратное утверждение неверно. Функция $d(n,S)$ вычислена лишь для некоторых значений аргументов (см. например [77],[88]) и до сих пор не решена, несмотря на элементарность формулировки, наиболее естественная задача о вычислении константы α такой, что

XLII

$$d(n,[2^{nH}]) \sim \alpha n. \tag{59}$$

XLIII

XLIV

Проблемы, аналогичные проблемам XVI и XXIII встают также и не находят решения, если ограничиться лишь изучением групповых кодов.

Понятие кодового расстояния может быть естественным образом введено и для простейших каналов с непрерывным множеством состояний. Там в качестве $\rho(.,.)$ надо брать обычное евклидово расстояние и исследование $d(n,s)$ сближается с известными геометрическими исследованиями о заполнении пространства сферами.

4.8. Предположение $S=[2^{nH}]$ введенное выше, является самым естественным, поскольку оно соответствует предположению о постоянстве скорости передачи. Однако некоторый интерес представляет исследование того, что происходит и при иной асимптотике S. Оказалось, в частности, что задача существенно упрощается, если предположить, что $S = const$. Здесь нашли решения аналоги проблем XLII (см.[2]; интересно, что при этом для бинарного канала $\alpha \geq \frac{1}{2}$, но $\alpha \to \frac{1}{2}$ при $S \to \infty$), проблем XLI, XLIII, XLIV, вопрос о вычислении констант a и b. Здесь, конечно, $\bar{b} > b$, так что случайный код оказывается хуже оптимального. Но уже для несимметрично бинарного канала без памяти (не говоря о более сложных случаях) асим-
XL птотика $e(Q^n, V^n, S)$ остается неисследованной.

Предположение $S = const$ означает, что рассматривается передача с очень малой скоростью. Другим крайним случаем является передача с максимально возможной скоростью, близкой к пропускной способности. М.Пинскером была указана следующая проблема. Пусть $N(n,\varepsilon)$ – наибольшее значение S такое, что сообщение $(p, \mathfrak{M}^S_\varepsilon)$ может быть передано через передающее устройство (Q^n, V^n). Надо изучить асимптотику $N(n,\varepsilon)$ при $\varepsilon = const$, $n \to \infty$. Легко видеть, что $e(Q^n, V^n, N(n,\varepsilon)) = \varepsilon$, так что изучение $\mathcal{N}(n,\varepsilon)$ есть одна из форм исследования асимптотики вероятности $e(Q^n, V^n, S)$. Нетрудно показать, что для однородного канала без памяти с симметрической матрицей,

$$\mathcal{N}(n,\varepsilon) \asymp n^{-1/2}\, 2^{n\bar{C} - \sigma u_\varepsilon \sqrt{n}} \tag{60}$$

где $\bar{C}$ - пропускная способность канала, u_ε - решение уравнения $\varphi(u_\varepsilon) = \varepsilon$, где $\varphi(.)$ - нормальная функция распределения, а $\sigma = \left[\frac{d\,m(o)}{d\,k}\right]^{1/2}$ (см. [48]). Некоторые оценки для $\mathcal{N}(n,\varepsilon)$ можно извлечь из работ Вольфовица

XLVI Интересно исследовать $\mathcal{N}(n,\varepsilon)$ для других каналов.

Все введенные выше постановки задачи об асимптотическом исследовании $e(Q^n, V^n, S)$ можно считать частными случаями следующей задачи: указать относительно простую функ-

XLVII цию $g(n,S)$ такую, чтобы при $n \to \infty$

$$e(Q^n, V^n, S) \asymp g(n, S)$$

равномерно по S (см. аналогичные задачи в предельной теории сумм случайных слагаемых ([52])). Однако эта проблема станет очевидной лишь после решения проблемы XXVII

4.9. Колмогоровым был указан несколько иной вариант постановок задач этого раздела. А именно: во изменение предположений раздела 4.1 будем считать, что $S = a^s$ где a и s - целые числа и отождествим значения сообщения с наборами $(E_{i_1}, \dots, E_{i_s})$, $i_k = 1, \dots, a$. Далее пространство $\tilde{X}$ значений сообщения на выходе будем считать совпадающим с X, а условия точности будем задавать набором s функций

$$\rho_k\big((E_{i_1}, \dots, E_{i_s}), (E_{j_1}, \dots, E_{j_s})\big) = \begin{cases} 0 & E_{i_k} = E_{j_k}, \\ 1 & E_{i_k} \neq E_{j_k}. \end{cases} \tag{61}$$

Наконец множество $\overline{W}$ будет состоять из одной точки $(0,\dots,0)$ (ср.разд. 1.6). Такое сообщение обозначим как $(p, \mathcal{N}^S)$. Через $g(Q,V,S)$ (ср.разд. 4.2) обозначим наименьшее ε такое, что сообщение $(p, \mathcal{N}^S_\varepsilon)$ может быть передано через передающее устройство (Q,V). С наглядной точки зрения можно сказать, что здесь сообщение на входе $\xi=(\xi_1,\dots,\xi_s)$ сообщение на выходе $\tilde{\xi}=(\tilde{\xi}_1,\dots,\tilde{\xi}_s)$ и

$$g(Q,V,S) = \inf \max_k \mathcal{P}\{\xi_k \neq \tilde{\xi}_k\}, \qquad (62)$$

где нижняя грань берется по всем способам передачи через передающее устройство (Q,V). Описанный критерий лучше отражает технические требования к точности передачи. Однако исследование величины $g(Q,V,S)$ существенно труднее, чем исследование $e(Q,V,S)$ и здесь пока что не получено

XLVIII никаких специфических результатов, хотя на этот случай могут быть переформулированы все вопросы, разбиравшиеся в параграфе 4.

5. <u>Единые способы передачи</u>. 5.1. Одно из основных предположений развитой выше теории состояло в том, что считались заданными распределение $P_\xi(\cdot)$ - сообщения на входе, и переходная функция $Q(\cdot,\cdot)$ передающего устройства. Однако во многих реальных ситуациях это предположение оказывается неразумным из-за того, что статистические параметры канала быстро меняются во времени, причем неоткуда взять априорные распределения этих параметров или же из-за того, что одна и та же приемно-передающая аппаратура должна быть приспособлена к работе в разнообразных условиях. Наконец, можно представить себе теоретико-игровую ситуацию, где один игрок выбирает распределения p и Q, а другой

выбирает способ передачи.

5.2. Математическое описание подобной ситуации сводится к следующему. Задано некоторое множество Γ значений параметра, причем каждому $\gamma \in \Gamma$ сопоставлено передающее устройство (Q_γ, V_γ) такое, что пространства сигналов на входе и выходе не зависят от γ. Далее задано множество сообщений (P_δ, W_δ), $\delta \in \Delta$ такое, что пространства значений сообщения на входе и выходе не зависят от δ. Мы будем говорить, что существует <u>единый способ передачи системы сообщений</u> $(P_\delta, W_\delta), \delta \in \Delta$ <u>через систему передающих устройств</u> $(Q_\gamma, V_\gamma), \gamma \in \Gamma$, если существуют кодирование $\mathcal{P}(\cdot,\cdot)$ и декодирование $\tilde{\mathcal{P}}(\cdot,\cdot)$ (не зависящие от γ и δ) такие, что для всех $\gamma \in \Gamma, \delta \in \Delta$ сообщение (P_δ, W_δ) может быть передано через передающее устройство (Q_γ, V_γ) при помощи кодирования $\mathcal{P}(\cdot,\cdot)$ и декодирования $\tilde{\mathcal{P}}(\cdot,\cdot)$.

<u>Назовем пропускной способностью системы</u> $(Q_\gamma, V_\gamma), \gamma \in \Gamma$

$$C(\Gamma) = \inf_{\eta} \sup_{\gamma \in \Gamma} \mathcal{I}(\eta, \eta_\gamma), \tag{63}$$

где (η, η_γ) связаны передающим устройством (Q_γ, V_γ), а нижняя грань берется по всем величинам η, для которых при всех γ существует такая пара (η, η_γ), и далее назовем <u>энтропией системы</u>

$$H(\Delta) = \sup_{\delta \in \Delta} H(P_\delta, W_\delta). \tag{64}$$

Из рассуждений раздела §4 легко следует, что <u>существование единого способа передачи влечет за собой неравенство</u>

$$H(\Delta) \leq C(\Gamma) \tag{65}$$

Понятие единого способа передачи было введено недавно в работах Блеквелла, Бреймана и Томасяна [9] и автора [22]. В работе [9] подробно изучен случай конечных однородных каналов без памяти. В работе [22] рассмотрен без доказательств и полных формулировок общий случай, когда Γ произвольно, а Δ состоит из одного элемента.

Отметим, что

$$C(\Gamma) \le \inf_{\gamma\in\Gamma} C(Q_\gamma, V_\gamma), \qquad (66)$$

причем равенство имеет место, когда система функций Q_γ является в некотором смысле (выпуклой). В общем случае в (64) может стоять знак $<$ (см. [22]). ~~[illegible]~~

5.3. Есть основания считать, что в достаточно широком классе случаев верна прямая теорема Шеннона для единых методов передачи, гласящая в общем случае, что при некоторых ограничениях на последовательность систем (Q_γ^t, V_γ^t), $\gamma\in\Gamma^t$ и (P_δ^t, W_δ^t), $\delta\in\Delta^t$ из того, что

$$\underline{\lim}_{t\to\infty} \frac{C(\Gamma^t)}{H(\Delta^t)} > 1 \qquad (67)$$

вытекает, что для любого $\varepsilon > 0$ существует столь большое T, что при всех $t \geqslant T$ существует единый способ передачи системы сообщений $(P_\delta^t, W_{\delta\varepsilon}^t)$, $\gamma\in\Gamma^t$ через систему передающих устройств $(Q_\gamma^t, V_{\gamma\varepsilon}^t)$. Вопрос об установлении достаточно общих (ср. 2.10) условий для выполнения
XLIX этой теоремы остается пока что открытым. Укажем однако те частные случаи, для которых нетрудно доказать ее верность. Теорема верна, если системы сообщений и передающих устройств удовлетворяют одному из следующих условий: 1) Γ^t состоит

из одного элемента (Q^t, V^t) и последовательность передающих устройств (Q^t, V^t) информационно-устойчива;

2) (Q^t_γ, V^t_γ), $\gamma \in \Gamma^t$ являются отрезками однородных каналов с дискретным временем, конечными множествами сигналов на входе и выходе и конечной памятью, причем переходные матрицы каналов таковы, что

$$\sum_{\tilde{y}_0 \in \tilde{Y}_0} q_0(y_0, f_0, \tilde{y}_0, f') \geq \alpha > 0 \tag{68}$$

для всех $y_0 \in Y_0$, $f_0 \in F$, $f' \in F$;

3) (Q^t_γ, V^t_γ), $\gamma \in \Gamma^t$ являются отрезками неоднородных каналов без памяти длины t, с конечными множествами сигналов, причем условия V^t_γ отсутствуют, а функции

$$Q^t_\gamma(y, \cdot) = Q_{a^1_\gamma}(y_1, \cdot)\, Q_{a^2_\gamma}(y_2, \cdot) \ldots Q_{a^t_\gamma}(y_t, \cdot) \tag{69}$$

где $(a^1_\gamma, a^2_\gamma, \ldots, a^t_\gamma)$, $\gamma \in \Gamma^t$ – это всевозможные наборы элементов $a \in A$ и $\{Q_a, a \in A\}$ – некоторое выпуклое множество переходных функций (Мы будем говорить, что множество переходных функций $\{Q_a, a \in A\}$ выпукло, если при Q_a, $a \in A$ и $Q_{\bar{a}}$, $\bar{a} \in A$, задаваемых матрицами (q^a_{ij}), $(q^{\bar{a}}_{ij})$, соответственно, функция, соответствующая матрице $(\lambda_i q^a_{ij} + (1-\lambda_i) q^{\bar{a}}_{ij})$, $0 \leq \lambda_i \leq 1$ также принадлежит A;

1') Δ^t состоит из конечного, независящего от t числа r сообщений (P^t_i, W^t_i), $i = 1, \ldots, r$ и все последовательности сообщений (P^t_i, W^t_i), $t = 1, 2, \ldots$ информационно устойчивы,

2') Условия W^t_δ не зависят от δ, а распределения P^t_δ можно описать как распределение набора $(f_\delta(\xi^\delta_1), \ldots, f_\delta(\xi^\delta_t))$, где ξ^δ_i образуют конечную однородную цепь Маркова с фиксированным числом состояний, причем все элементы матриц вероятностей перехода равномерно по δ больше константы $\alpha>0$.

3) Условия W_δ^t не зависят от δ, а сообщения (P_δ^t, W_δ^t) являются неоднородными по времени сообщениями с покомпонентными условиями точности воспроизведения и независимыми компонентами, причем $P_\delta^t = P_{a_\delta^1}\times\ldots\times P_{a_\delta^t}$, где $(a_\delta^1,\ldots,a_\delta^t), \delta\in\Delta$ это всевозможные наборы элементов $a\in A$ и $\{P_a, a\in A\}$ выпуклое множество распределений вероятностей на конечном пространстве X_0.

Приведенные примеры показывают, что прямая теорема Шеннона верна в достаточно общей обстановке. С другой стороны общность ее все же меньше, чем общность теоремы Шеннона раздела 2.10. Так, сформулированное утверждение будет, например, вообще говоря неверно, если среди (P_δ, W_δ) есть сообщения с одинаковыми P_δ, но разными W_δ (здесь все сведется к необходимости удовлетворить более сильное условие точности, являющееся пересечением нескольких W_δ, а для решения такой задачи ответ дается методами раздела 2). Далее, если в примере 3 (аналогично 3') множество $\{Q_a, a\in A\}$ невыпукло, то для того, чтобы теорема Шеннона осталась верна нужно в определении (63) (аналогично (64)) заменить систему Q^γ на ее выпуклую оболочку.

5.4. Отметим некоторые важные частные случаи формул (63) и (64) для пропускной способности и энтропии систем. Пусть $(Q_\gamma, V_\gamma), \gamma\in\Gamma$ система всех каналов без памяти длины n с аддитивным шумом, мощностью шума P_n и мощностью сигнала P_s. Тогда $C(\Gamma)$ задается формулой Шеннона (32). Далее, если $(Q_\gamma, V_\gamma), \gamma\in\Gamma$ система всех передающих устройств с конечной полосой пропускания W с аддитивным шумом для передачи за время T, $C(\Gamma)$ - задается асимптотически формулой Шеннона (35). Наконец, если (P_δ, W_δ), $\delta \in \Delta$ - система всех сообщений с покомпонентными условиями точности, задающими ограничением среднеквадратичной ошибки константой P_d и дисперсиями компонент сообщения на входе, равными P_c, то для $H(\Delta)$ верно выражение (34). Формулы (32), (34), в особенности (35), выводимые обычно для гауссовского случая, применяются на практике в много более широком классе случаев, в том числе и тогда, когда о характере распределений ничего определенного сказать нельзя. Оправданием подобной практике может служить приводимая выше новая интрепретация этих формул.

I 5.5. Отметим, что остается открытым вопрос о создании общих методов для вычисления пропускной способности $C(\Gamma)$ и энтропии $H(\Delta)$, а также соответствующих "оптимальных стратегий" (т.е. значений ξ, γ и δ, на которых достигаются экстремумы в (62) и (63) (ср.пробл. XУ1, XУП, XУШ). Далее кажется интересным исследовать в стиле задач, разобранных в разделе 4, асимптотику вероятностей ошибок при единых методах передачи.

II В качестве первого шага можно предложить исследование вероятности ошибки для системы всех каналов без памяти, матрицы которых являются квадратными фиксированного порядка, причем их диагональные элементы $p_{ii} \geq 1-\varepsilon$, где ε - заданная константа.

6. Кодирование с использованием дополнительных сведений. 6.1. Иногда естественно предположить, что на входе

канала известны и могут быть поэтому использованы при кодировании некоторые дополнительные сведения о значении состояния канала или о значении сигналов на выходе в предыдущие моменты времени (подобные сведения может принести канал обратной связи или некоторые методы зондирования среды, через которую ведется передача). Встает вопрос о том, как при этом изменится постановка и решение проблемы Шеннона. Для некоторых частных случаев этот вопрос изучался в работах Шеннона ([81], [84]) и автора [18].

6.2. Для простоты мы дадим полную математическую формулировку задачи лишь для случая однородноро канала с дискретным временем, хотя читатель без труда поймет как это определение распространяется на общий случай. Мы предположим, что кроме об"ектов, входящих в определение отрезка длины n однородного канала с дискретным временем, задана последовательность измеримых пространств $(\mathcal{D}_k, S_{\mathcal{D}}^k)$, называемых <u>пространствами дополнительных сведений в момент</u> k. Кроме того задана переходная функция

$$R_k(y_1,\dots,y_{k-1},\tilde{y}_1,\dots,\tilde{y}_{k-1},f_1,\dots,f_k,A),\ k=1,\dots,n$$
$$y_i\in Y_0,\ \tilde{y}_i\in\tilde{Y}_0,\ f_i\in F,\ A\in S_{\mathcal{D}}^k$$

измеримая по совокупности первых $3k-2$ переменных и являющаяся мерой по A, С наглядной точки зрения $R_k(\cdot,\cdot)$ задает распределение вероятностей для сведений, наличных на входе канала в момент k, если заданы значения сигналов и состояний канала за предыдущие время. Таким образом, приводимая формулировка включает в себя и тот случай, когда сведения могут быть случайными, например, из-за помех в канале обратной связи. Однако при этом предполагается, что фиксирован и неизменен метод передачи по каналу обратной

LII связи. Интересной казалась бы задача, в которой считается заданным лишь сам канал обратной связи, а метод передачи по нему выбирается оптимальным. Подходов к решению подобной задачи не видно.

Укажем важнейшие частные случаи. Если $\mathcal{D}_k = \mathcal{Y}_0^{k-1} \times \tilde{\mathcal{Y}}_0^{k-1} \times F^k$ и $R_k(d, \cdot)$, $d \in \mathcal{D}_k$ сосредоточено в точке d, то мы будем говорить, что задан <u>канал с полными сведениями о прошлом.</u> Если $\mathcal{D}_k = \tilde{\mathcal{Y}}_0^{k-1}$ и $R_k(\cdot,\cdot)$ зависит лишь от $\tilde{y}_1,\dots,\tilde{y}_{k-1}$ и сосредоточено в соответствующей точке, то мы будем говорить, что задан <u>канал с полной обратной связью</u> (иногда, не совсем точно говорят, что здесь имеется канал обратной связи с бесконечной пропускной способностью). Если, наконец, $\mathcal{D}_k = F$, $R_k(\cdot,\cdot)$ зависит лишь от f_k и, как мера, сосредоточена в соответствующей точке, то мы будем говорить, что задан <u>канал с полными сведениями о его состоянии</u>.

6.3. Мы будем говорить, что задан <u>способ передачи сообщения (p, W) с использованием дополнительных сведений,</u> если указана система величин $\xi, \tilde{\xi}, \eta_1,\dots,\eta_n, \tilde{\eta}_1,\dots,\tilde{\eta}_n$, $\varphi_1,\dots,\varphi_n, \delta_1,\dots,\delta_n$ со значениями в $X, \tilde{X}, \mathcal{Y}_0, \tilde{\mathcal{Y}}_0, F, \mathcal{D}$ соответственно такая, что 1) пара $(\xi, \tilde{\xi})$ образуют сообщение (p, W), 2) при всех k и всех $A \in S_{\tilde{\mathcal{Y}}} \times S_F$ условная вероятность

$$\mathcal{P}\{(\tilde{\eta}_k, \varphi_k) \in A / \eta_1,\dots,\eta_k, \tilde{\eta}_1,\dots,\tilde{\eta}_{k-1}, \varphi_1,\dots,\varphi_{k-1}, \xi\} =$$

$$= Q_0(\eta_{k-1}, \varphi_{k-1}, A) \qquad (70)$$

3) распределение величины φ_0 совпадает с начальным распределением канала P_0, 4) при любом k и любом $A \in S_{\mathcal{D}}^k$

$$\mathcal{P}\{\delta_k \in A / \eta_1,\dots,\eta_{k-1}, \tilde{\eta}_1,\dots,\tilde{\eta}_{k-1}, \tilde{\varphi}_1,\dots,\tilde{\varphi}_{k-1}, \delta_1,\dots,\delta_{k-1}\} = \qquad (71)$$

$$= R_k(\eta_1,\dots,\eta_{k-1}, \tilde{\eta}_1,\dots,\tilde{\eta}_{k-1}, \tilde{\varphi}_1,\dots,\tilde{\varphi}_{k-1}, A).$$

5) при любом k и любом $A \in S_y$

$$\mathcal{P}\{\eta_k \in A / \eta_1,\dots,\eta_{k-1},\tilde{\eta}_1,\dots,\tilde{\eta}_{k-1},\tilde{\varphi}_1,\dots,\tilde{\varphi}_k,\delta_1,\dots,\delta_k,\xi\} = \mathcal{P}\{\eta_k \in A / \eta_1,\dots,\eta_{k-1},\delta_k,\xi\}, \tag{72}$$

6) при любом $A \in S_{\tilde{X}}$

$$\mathcal{P}\{\tilde{\xi} \in A / \eta_1,\dots,\eta_n,\tilde{\eta}_1,\dots,\tilde{\eta}_n,\varphi_1,\dots,\varphi_n,\delta_1,\dots,\delta_n,\xi\} = \mathcal{P}\{\tilde{\xi} \in A / \tilde{\eta}_1,\dots,\tilde{\eta}_n\} \tag{73}$$

Наглядный смысл условий 1), 2), 3), 4) ясен. Условие 5) означает, что при выборе сигнала на входе в k -ый момент используется лишь знание сообщения на входе, значений сигнала на входе в предыдущие моменты и дополнительные сведения δ_k. Условие 6) означает, что декодирование основывается лишь на знании сигнала на выходе.

Если пространства дополнительных сведений $\mathcal{D}_k$ состоят каждое из одного элемента, то это означает, что дополнительных сведений нет. Нетрудно понять, что в этом случае введенное определение эквивалентно определению раздела 1.10.

6.4. Шеннон рассмотрев в [84] случай канала с полными сведениями о состоянии такого, что $q_0(y_0,\ell,\tilde{y}_0,\ell')$ зависит лишь от ℓ' (если дополнительных сведений нет, то этот канал превращается в канал без памяти). Он указал прием, позволяющий сводить исследование канала с дополнительными сведениями к обычному каналу. Его рассуждения легко распространяются на самый общий случай. А именно: всегда можно построить передающее устройство такое, что существование способа передачи сообщения (p,W) через это устройство эквивалентно существованию способа передачи того же сообщения через первоначальный канал с использованием дополнительных сведений.

Строится это передающее устройство так: пространство $\tilde{\mathcal{Y}}$ совпадает также, как и в первоначальном канале, с $\mathcal{Y}_0^n$: Далее вводятся пространства $\mathcal{Y}_0(\mathcal{D}_k)$, $k=1,\dots,n$ всех измеримых отображений $y(d_k)$ из $\mathcal{D}_k$ в $\mathcal{Y}$ с естественным образом вводимой σ-алгеброй измеримых подмножеств и полагается $\mathcal{Y} = \mathcal{Y}_0(B_1) \times \dots \times \mathcal{Y}_0(B_n)$. Что касается переходной функции нового передающего устройства, то в случае канала с полными сведениями о состоянии надо считать передающее устройство каналом с тем же, что и раньше пространством состояний и новой переходной плотностью.

$$\tilde{q}_0(y(d_k), f_0, \tilde{y}_0, f') = q_0(y(f_0), f_0, \tilde{y}_0, f'),$$

где $y(f_0)$ значение функции $y(d_k)$ при $d_k = f_0$

В общем случае определение переходной функции громоздко и мы укажем лишь, что его основная идея остается той же самой: в тех случаях, когда на вход передающего устройства подано $y(d_k) \in \mathcal{Y}(\mathcal{D}_k)$ и сведения о прошлом (которые могут быть математически определены и для нового передающего устройства, но теперь не должно влиять на кодирование) оказались равными d_k^0 надо, чтобы передача происходила так, как и в прежнем канале, на вход которого подан сигнал $y(d_k^0)$. Таким образом, по крайней мере формально, проблема Шеннона для канала с дополнительными сведениями сводится к проблеме Шеннона в формулировке раздела 1. Однако, поскольку построенное передающее устройство оказывается очень сложным, то к нему неприменимы критерии информационной устойчивости, обсуждавшиеся в § 2, и потому остается открытым вопрос о выводе специфических для этого случая условий информационной устойчивости.

LIII

6.5. Самым интересным является вопрос о вычислении пропускной способности каналов с дополнительными сведениями. Отметим прежде всего следующий факт. Пропускная способность канала с полными сведениями о прошлом всегда совпадает с пропускной способностью (того же) канала с полными сведениями о его состоянии. Этот факт является следствием более общего результата, который мы не будем приводить здесь подробно и который с наглядной точки зрения показывает, что обратная связь полезна для увеличения пропускной способности лишь в той мере, в какой она доставляет сведения о состоянии канала. Далее, используя метод, развитый в [18], нетрудно показать, что пропускная способность канала с дополнительными сведениями совпадает с обычной пропускной способностью этого канала, если выполнено следующее условие: Пусть $\eta, \tilde{\eta}_\ell$ случайные величины, связанные передающим устройством с пространствами сигналов на входе и выходе $(Y_0, S_{Y_0}), (\tilde{Y}_0, S_{\tilde{Y}_0})$ и переходной функцией $Q_0(\cdot, \ell, \cdot)$, тогда $\mathcal{I}(\eta, \tilde{\eta}_\ell)$ не зависит от ℓ, а зависит лишь от распределения η. Применяя высказанное утверждение к каналу без памяти, получаем результат Шеннона [81], гласящий, что обратная связь не увеличивает пропускной способности канала без памяти (В момент опубликования работы [18], где доказан тот же результат, автор не был, к сожалению, знаком с работой Шеннона).

Другой важный пример канала, обладающего указанным свойством, дают нам гауссовские передающие устройства (если превратить их в канал приемом, описанным в разделе 1). Дело в том, что для таких каналов $Q_0(\cdot, \ell, \cdot)$ ~~является~~ задает гауссовское передающее устройство, причем от ℓ зависит лишь средние значения, а не вторые моменты. Изменение же среднего значения η_ℓ не влияет на информацию $J(\eta, \eta_\ell)$. Таким образом обратная связь не увеличивает пропускной способности гауссовского передающего устройства. Это было впервые отмечено М.Пинскером. Шенноном [84] вычислена пропускная способность для упомянутого выше класса каналов с полными сведениями о состоянии. Другим интересным примером могут служить каналы со случайным параметром и полными сведениями о значении параметра $\beta \in B$. Здесь (ср. (36)), повидимому, средняя пропускная способность.

$$\bar{C} = \int_B \bar{C}_\beta \, P(d\beta) \tag{74}$$

где $\bar{C}_\beta$ - средняя пропускная способность для условного канала с параметром β. Эта формула легко доказывается (ср. [60]) для случая, когда пространство B -конечно. Поскольку она находит интересные приложения (с ее использованием Овсеевичем, Пинскером и Цыбаковым показано, что в некоторых физически реальных ситуациях обратная связь может в полтора раза увеличить пропускную способность канала) в случае непрерывного распределения параметра, интересно доказать ее и в общем случае (ср. разд. 3.6). Важно научиться вычислять пропускную способность каналов с дополнительными сведениями и в других случаях.

LIV

LV

Интересно также исследовать вопрос о том, может ли для однородных каналов без памяти использование обратной связи уменьшить оптимальную вероятность ошибки $e(Q^n, V^n, [2^{nH}])$ (см. 4.2). Нетрудно показать, что для каналов с симметричными матрицами при $H \geq H_{crit}$ использование обратной связи

не меняет констант a и b в соотношении (45). В работе Элайаса [27] содержится замечание, которое может быть интрепретировано как утверждение о том, что при $H < H_{crit}$ для бинарного симметричного канала с обратной связью оказывается верным соотношение (45) с константами a и b, задаваемыми снова равенством (47). Автору неизвестно доказательство этой теоремы,
LVI а также то, в какой мере она распространяется на произвольный канал с симметричной матрицей (Как указано в [27], аналогичный факт легко проверяется для симметричного бинарного канала со стиранием). Очередной кажется также задача исследовать
LVII довать $e(Q^n, V^n, [2^{nH}])$ для гауссовских каналов без памяти с аддитивным шумом и обратной связью.

7. <u>Новые применения понятий Шенноновской теории информации.</u> 7.1. В предыдущих разделах этого обзора мы рассматривали лишь вопросы, входящие в рамки сформулированной в разделе 1 основной проблемы Шеннона об оптимальной передаче информации. Здесь мы хотим кратко обсудить иные, пока что лишь едва наметившиеся, направления приложений понятий энтропии и информации. Не исключено, что в будущем все эти направления сольются в некую единую теорию, но даже контуры такой единой теории пока что не видны.

7.2. Первым из таких направлений является использование обобщенной энтропии пары распределений как меры их различия в задачах математической статистики. Обзор многочисленных работ на эту тему содержится в недавней книге Калбекаа [51]. В основном опубликованные работы содержат перечисление свойств обобщенной энтропии, подтверждающих, что ее удобно использовать в качестве меры статистического различия. Более важным кажется нам попытка показать, что через обобщенную энтропию дается асимптотический ответ на некоторые классические проблемы математической статистики. Для независимых наблюдений это сделано в работах Мурье [53] и Айвазяна [1]. В новой работе Добрушина, Пинскера и Ширяева указывается, что отмеченные асимптотические результаты распространимы на очень широкий класс зависимых испытаний.

7.3. Второе многообещающее направление представлено пока что лишь в форме исследования с использованием понятия энтропии известной из популярных книг по математике задачи о выделении фальшивой монеты наименьшим числом взвешиваний. Такое исследование начато в [49] и подробно развито в книге [45]. Здесь можно сформулировать общую проблему о минимальном числе экспериментов, необходимых для достижения какого-то знания и, повидимому, асимптотическое решение этой проблемы может быть при некоторых условиях дано через понятие энтропии.

7.4. Автору непонятны причины, по которым энтропия и информация появились в некоторых теоретико-игровых конструкциях (см. [50]). Перспективы этого направления неясны.

Автор благодарит А.Н.Колмогорова, М.С.Пинскера и В.М. Яглома. Постоянный контакт с ними во многом отразился на содержании этой статьи.

Р. Добрушин

ЛИТЕРАТУРА

1. Айвазян С.А. Сравнение оптимальных свойств критериев Неймэна-Пирсона и Вальда. Теор. вер. и ее прим., IУ, 1959, 86-93.

2. Бакут Л.А. - К теории корректирующих кодов с произвольным основанием, научн. докл. Высш. школы, Радиот.и электрон., № 1, 1959.

3. Бахвалов Н.С.- К вопросу о числе арифметических действий при решении уравнения Пуассона для квадрата методом конечных разностей. Докл. АН СССР, 113, 1957, 252-254.

4. Башарин Г.П. - О статистической оценке энтропии последовательности независимых случайных величин. Теор.вер. и ее прим., IУ, 1959, 361-364.

5. Bellman P., Dynamic programming, Princeton, 1957, Princeton University Press

6. Blackwell D., The Entropy of function of finite state Markov chains, Trans. of the first Prague conf. in Inform. Theory, Stat. des. funct., random proc., Prague, 1957, 13-20.

7. Blackwell D., Infinite codes for memoryless channels, Ann. Math. Stat., 30, 1959, 1242-1244

8. Blackwell D., Breiman L., Thomasian A.J., Proof of Shannon's transmission theorem for finite-state indecomposable channels, Ann. Math. Stat., 29, 1958, 1209-1220.

9. Blackwell D, Breiman L., Thomasian A.J., The capacity of a class of channels, Ann. Math. Stat., 30, 1959, 1229-1241.

10. Бородин Л.Ф. - Некоторые вопросы теории построения корректирующих кодов, научн.-техн.об. Радиотехн. и электр. им.А.С.Попова. Сб. трудов, вып. 2, М., 1958, 110-151.

11. Breiman L., The individual ergodic theorem of information theory, Ann. Math. Stat., 29, 1957, 809-811.

12. Carleson L., Two remarks on the basic theorems of information theory, Scand. Math. 6, 1958, 175-180

13. Хинчин А.Я. - Понятие энтропии в теории вероятностей. Усп. мат.наук, 8, № 3, 1953, 3-20.

14. Хинчин А.Я. - Об основных теоремах теории информации. Усп. мат.наук, 11, № 1, 1956, 17-75.

15. Cramer H. Sur un nouveau théorem-limite de la théorie des probabilités, Act. Sc. et. ind., Paris, 1938

16. Davenport W.B, Root W.L., An Introduction to the theory of random signals and noise, N-Y, Toronto, Lnd., Mc Graw-Hill Book Company Inc, 1958

17. Добрушин Р.Л. - По поводу формулировки основной теоремы Шеннона (Резюме доклада на заседании научн.-исслед. семинара по теории вероятностей 19/Ш 1957 г.). Теор.вер. и ее прим., 2, 1957, 480-482.

18. Добрушин Р.Л. - Передача информации по каналу с обратной связью. Теор.вер. и ее прил., 3, 1958, 395-412.

19. Добрушин Р.Л. - Упрощенный метод экспериментальной оценки энтропии стационарной последовательности. Теор. вер. и ее прил., 3, 1958, 462-464.

20. Добрушин Р.Л. - Общая формулировка основной теоремы Шеннона в теории информации. Докл. АН СССР, 126, 1959, 474-477.

21. Дубрушин Р.Л. - Общая формулировка основной теоремы Шеннона в теории информации. Усп.мат. наук, 14, № 6, 1959, 3-104.

22. Добрушин Р.Л. - Оптимальная передача информации по каналу с неизвестными параметрами. Радиот. и электрон., 4, 1959, 1951-1956.

23. Добрушин Р.Л. - Предельный переход под знаком информации и энтропии. Теор.вер. и ее прил., 5, 1960, 28-37.

24. Doob J.L., Editorial, IRE Trans. Inf. Theory, IT-5, № 1, 1959, 3

25. Дынкин Е.Б. - Основания теории марковских процессов, М., 1959, Физматгиз.

26. Elias P., Coding for ~~two~~ noisy channels, IRE Conv. Rec., 3, № 4, 1955, 37-46.

27. Elias P, Coding for two noisy channels, Inform. Theory 3rd London Sympos., 1955, Lnd, 1956, 61-74.

28. Epstein M.A, Algebraic decoding for a binary erasure channels, IRE Nat. Conv. Rec, 1958, 6, №4, 56-59

29. Ерохин В. - ε-энтропия дискретного случайного об"екта. Теор. вер. и ее прил., 3, 1958, 103-107.

30. Фаддеев Д.К. - К понятию энтропии конечной вероятноостной схемы. Усп. мат.наук, 11, № 1, 1958, 227-231.

31. Feinstein A., A new basic theorem of information theory, Trans. IRE, 1954, PG-IT-4, 2-22

32. Feinstein A, Error bounds in noisy channels without memory, IRE Trans. Inform. Theory, 1955, 1, №2, 13-14

33. Feinstein A, Foundations of information theory, N-Y, Mc Grow-Hill, 1958

34. Feinstein A, On the coding theorem and its converse for finite-memory channels, Inf. and Contr. 1959, 2, [illegible], 25-44

35. Feinstein A., On the coding theorem and its converse for finite-memory channels, Nuovo Cimento, Supplem. XIII, №2, 1959, 345-352

36. Fleischer I., The central concepts of ~~information~~ communication theory for infinite alphabets, J. Math. and Phys, 37, 1958, 223-228

37. Fontaine A.B, Peterson W.W., On coding for the binary symmetric channels, Trans. Amer. Inst. Electr. Engr., I, 77, 1958, 638-647.

38. Гельфанд И.М., Колмогоров А.Н., Яглом И.М. К общему определению количества информации. Докл. АН СССР, 111, 1956, 745-748.

39. Гельфанд И.М., Колмогоров А.Н., Яглом И.М. - Количество информации и энтропия для непрерывных распределений. Труды третьего матем. с"езда, том Ш, М., Из-во АН СССР, 1958, 300-320.

40. Гельфанд И.М., Яглом А.М. - О вычислении количества информации о случайной функции, содержащейся в другой такой функции. Усп. мат.наук, 12, № 1, 1957, 3-52.

41. Гнеденко Б.В., Колмогоров А.Н. - Предельные распределения для сумм независимых величин. М.-Л. Гостехиздат, 1958.

42. Gilbert E., N, Moor E.F., A variable-length Binary Encodings, Bell Syst. Techn. J., 38, 1959, 933-967.

43. Good J.J, Doog K. Caj, A paradox concerning rate of information, Inf. and Contr., 1, 1958, 91-112

44. Hamming R.W., Error detecting and Error correcting codes, Bell. Syst. Techn. J., 29, 1950, 147-160

45. Яглом А.М., Яглом И.М. - Вероятность и информация, изд. -ое, М., Гостехиздат, 1960.

46. Jakobs K., Die Übertragung diskreter Informationen durch periodische und fastperiodische Kanäle Math. Ann., 137, 1959, 125-135

47. Joshi A.D, A note on upper bounds for minimum distance codes, Inf. and contr., 1, 1958, 289--295.

48. Юшкевич А.А. - О предельных теоремах, связанных с понятием энтропии цепей Маркова. Усп. мат.наук, 8, 1953, 177-180.

49. Kellog P.J., Kellog D.J, Entropy of information and the odd ball problem, Journ. of Appl. Phys., 25, 1954, 1438-1439.

50. Kelly J.L. Jr, A new interpretation of Information rate, Bell Syst. Techn. J, 35, 1956, 917-926.

51. Kullback S., Information theory and statistics, John Willey and Sons, Inc., N-Y, Champan and Hall, Ltd, Lnd, 1959

52. Колмогоров А.Н. - Некоторые работы последних лет в области предельных теорем теории вероятностей. Вестн. МГУ, № 10, 1953, 29-38.

53. Колмогоров А.Н. - Теория передачи информации. Сессия АН СССР по научн.пробл. автомат.производства, 1956, Пленарные заседания, М. Изд. АН СССР, 1957, 66-99

54. Колмогоров А.Н. - Новый метрический инвариант динамических систем и автоморфизмов пространств Лебега. Докл. АН СССР, 119, 1958, 861-864.

55. Колмогоров А.Н. - Об энтропии на единицу времени, как метрическом инварианте автоморфизмов. Докл. АН СССР, 124, 1959, 754-755.

56. Колмогоров А.Н., Тихомиров В.Н. ε -энтропия и ε -емкость множеств в метрических пространствах. Усп. мат., наук, 14, № 2, 1959, 3-86.

57. Lomnizky Z.A, Zaremba S.K, The asymptotic distribution of the amount of transmitted information, Inf. and Contr., 2, 1959, 266-284

58. Mc Cluskey E.J, Jr, Error correcting codes - a linear programming approach, Bell Syst. Techn. J., 38, 1959, 1485-1512.

60. Mc Millan B., The basic theorem of information theory, Ann. Math. Stat., 24, 1953, 196-219.

61. Muroga S, On the capacity of a discrete channel. 1. Mathematical expression of capacity of a channel which is disturbed by noise in its very one symbol and expressible in one state diagram., J. Phys Soc. Japan, 8, 1953, 484-494

62. Muroga S., On the capacity of a noisy continuous channel, IRE Trans. Inf. Theory, IT-3, 1957, 44-51.

63. Mourier E., Étude du choix entre-deux lois de probabilité, C.R. Acad. Sci. Paris., 223, 1946, 712-714

64. Nedoma J., The capacity of a discrete channel, Trans. of the first Prague Conf. in Inform. Theory, stat. des. funct., random proc., Prague, 1957, 143-182

65. Овсеевич И.А., Пинскер М.С. - Скорость передачи информации, пропускная способность многоканальной системы и прием по методу линейно-операторного преобразования. Радиотехн., 1959, № 3, 9-21.

66. Perez A, Notions généralisées d'incertitude, d'entropie et d'information du point de vue de la theorie de martingales, Trans. of the first Prague conf. in Inform. Theory, Stat. des. funct., random proc., Prague, 1957, 183-208

67. Perez A, Sur la théorie de l'information dans le case d'un alphabet abstrait, там же, 209-244

68. Perez A, Sur la convergence des incertitudes, entropies et information enchantillon vers leurs valeurs vraises, там же, 245-252

69. Перез А. - Теория информации с абстрактным алфавитом. Обобщенные виды предельной теоремы Мак-Миллана для случая дискретного и непрерывного времени. Теор. вер. и ее прим., 4, 1959, 105-109.

70. Пинскер М.С. Количество информации о гауссовском случайном процессе, содержащееся во втором процессе, стационарно с ним связанном. Докл. АН СССР, 99, 1954, 213-216.

71. Пинскер М.С. Вычисление скорости создания сообщений стационарным случайным процессом и пропускной способности стационарного канала Докл. АН СССР, 111, 1956, 753-756.

72. Пинскер М.С. Экстраполирование однородных случайных полей и количество информации о гауссовском случайном поле, содержащейся в другом гауссовском случайном поле. Докл. АН СССР, 112, 1957, 815-818.

73. Пинскер М.С. Экстраполирование случайных векторных процессов и количество информации, содержащейся в одном векторном стационарном случайном процессе, относительно другого, стационарно с ним связанного. Докл. АН СССР, 121, 1958, 49-51.

74. Powers K, A prediction theory approach to information rates, IRE conv. Rec., 4, №4, 1956, 132-139

75. Reniy A., Balatoni J., Über den Begriff der Entropie, Math-Forschungsber., 4, 1957, 117-134.

76. Розенблат-Рот М. - Энтропия стохастических процессов. Докл. АН СССР, 112, 1957, 16-19.

77. Розенблат-Рот М. - Теория передачи информации через статистические каналы связи. Докл. АН СССР, 112, 1957, 202-205.

78. Розенблат-Рот М. - Нормированная ε -энтропия множеств и передача информации непрерывных источников через непрерывные каналы связи. Докл. АН СССР, 130, 1960, 265-268.

79. Рубинштейн Г.Ш., Урбаник К. - Решение одной экстремальной задачи. Теор. вер. и ее прим. 2, 1957, 375-377.

80. Shannon C.E., A mathematical theory of communication, Bell Syst. Techn. J., 27, 1948, 379-423, 623-656

81. Shannon C.E., The zero capacity of a noisy channel, IRE Trans Inf. Theory, IT-2, 1956, 8-19

82. Shannon C.E., Some geometric results in channel capacity, Nachr. Techn. Fachber., 6, 1956, 13-15

83. Shannon C.E., Certain results in coding theory for noisy channels, Inf. and Contr., 1, 1957, 6-25

84. Shannon C.E., Channels with side information at the transmitter, IBM J. Res. and Devel., 2, 1958, 289-293

85. Shannon C.E., Coding theorems for a discrete source with a fidelity criterion, IRE Nat. Conv. Rec., 7, No4, 1959, 142-163

86. Shannon C.E., Probability of error for optimal codes in a Gaussian channel, Bell Syst. Techn. J., 38, 1959, 611-655

87. Shannon C.E., McCarthy J., Automata studies, Princeton, Princeton University Press, 1956

88. Shapiro H.S, Slotnick D.L., On the mathematical theory of error-correcting codes, IBM. J. Res. Develop., 3, 1959, 25-34

89. Slepian D., A class of binary signalling alphabets, Bell Syst. Techn. J., 35, 1956, 203-234.

90. Takano K., On the basic theorem of information theory, Ann. Inst. Math. Stat. Tokyo, 9, 1958, 53-77

91. Ulrich W., Non-binary error-correcting codes, Bell Syst. Techn. J., 36, 1957, 53-77

92. Витушкин А.Г. - Оценка сложности задач табулирования. Физматгиз, М., 1959.

93. Wolfowitz J., The coding of message subject to chance errors., Ill. J. of Math., 1, 1957, 591-606.

94. Wolfowitz J. An upper bound of the rate of transmission of message, Ill. J. of Math., 2, 1958, 137-141.

95. Wolfowitz J. The maximum achievable length of an error-correcting code, Ill. J. of Math., 2, 1958, 454-458.

96. Wolfowitz J., Information theory for mathematics, Ann. Math. Stat., 29, 1958, 351-356

97. Wolfowitz J., Strong converse of the coding theorem for semicontinuous channels, Ill. J. of Math., 3, 1959, 477-489

98. ~~Wolfowitz~~ Wozencraft J.M. Sequential decoding for reliable communication, IRE Nat conv. Rec., 5, №2, 1957, 11-25

99. Цареградский И.П. - Замечание о пропускной способности стационарного канала с конечной памятью. Теор. вер. и ее прим., 3, 1958, 84-96.

100. Цзян-Цзе-Пей - Замечание об определении количества информации. Теор. вер. и ее прим., 3, 1958, 99-103.

101. Цыбаков Б.С. - О пропускной способности двулучевых каналов связи. Радиотехн. и электр., 4, 1959, 1117-1123.

102. Цыбаков Б.С. - О пропускной способности каналов и большим числом лучей. Радиотехн. и электр., 4, 1959, 1427-1433.

103. Цыбаков Б.С. - Пропускная способность некоторых многолучевых каналов. Радиотехн. и электр., 4, 1959, 1602-1608.

On Applicability of the Central Limiting Theorem by Yu A Rosanov

Ю.А. РОЗАНОВ.

О применимости центральной предельной теоремы.

Введение.

Предельные теоремы, установленные для классического случая сумм независимых величин, не могли удовлетворить тех запросов, которые возникали как в самой теории вероятностей, так и в её приложениях.

Ещё начиная с работ Бернштейна[1], были предприняты попытки распространить эти теоремы на случай зависимых величин. Наиболее окончательные результаты в этом направлении, как известно, были получены для величин, связанных в цепь Маркова[2]-[6].

Совершенно ясно, что сколь-нибудь общие предельные теоремы можно было надеяться получить лишь для величин, в каком-то смысле слабо зависимых. Естественным образом возникло понятие m-зависимых случайных величин, на которые без особого труда обобщаются результаты классического случая независимых величин[6]-[8].

В последнее время появилось несколько различных условий слабой зависимости, использование которых привело к установлению ряда новых предельных теорем[9]-[11].

Наиболее широкое из них было сформулировано Розенблатом[9] для случая стационарной последовательности $\xi(t)$ и заключается в следующем:

$$|P(AB) - P(A)P(B)| \leq \alpha(\tau) \qquad (1)$$

где $A \in \mathfrak{M}^{t}_{-\infty}$, $B \in \mathfrak{M}^{\infty}_{t+\tau}$ / $\mathfrak{M}^{t}_{s}$ есть σ-алгебра, порождённая событиями вида $\{\xi(u) < x\}$, $s \leq u \leq t$ / и $\alpha(\tau) \to 0$ при $\tau \to \infty$.

В этой форме условие (1), которое, следуя Розеблатту,[9] будем называть условием сильного перемешивания, применимо к произвольному случайному процессу $\xi(t)$ / и вообще, к некоторому семейству σ-алгебр $\mathfrak{M}_s^t$ /.

Условие (1) сильного перемешивания выполняется для широкого класса эргодических марковских, а также гауссовских процессов. Было установлено[12], что для стационарного гауссовского процесса свойство сильного перемешивания связано с гладкостью его спектральной плотности; например, в случае дискретного времени оно всегда выполняется, если спектральная плотность непрерывна и не обращается в нуль.

Для величин $\xi(t)$, удовлетворяющих условию (1) сильного перемешивания, наряду с самой центральной предельной теоремой[11] были получены её уточнения[17] / асимптотические разложения и большие уклонения /.

Другим, более ограничительным, чем условие (1), является условие, использованное Ибрагимовым[10] / также в случае стационарной последовательности /:

с вероятностью I

$$\sup_{A \in \mathfrak{M}_{t+\tau}^{\infty}} |P(A/\mathfrak{M}_{-\infty}^{t}) - P(A)| \leq \beta(\tau) \downarrow 0 \quad (2)$$

при $\tau \to \infty$.

Использование этого условия позволило получить заново предельные теоремы для марковских процессов [4], а также для некоторых специальных процессов, имеющих интерес в теории чисел [15].

§ I. ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА ДЛЯ АДДИТИВНЫХ СЛУЧАЙНЫХ ФУНКЦИЙ.

Пусть $\mathcal{H}^{s,t}(\omega)$ есть семейство случайных величин, аддитивно зависящих от интервала (s,t), т.е.

$$\mathcal{H}^{s,u}(\omega) + \mathcal{H}^{u,t}(\omega) = \mathcal{H}^{s,t}(\omega) \tag{3}$$

с вероятностью 1 для всех $s \le u \le t$, и пусть

$$m(s,t) = M\mathcal{H}^{s,t}, \quad \sigma^2(s,t) = D\mathcal{H}^{s,t}.$$

Важнейшими примерами таких величин являются величины вида

$$\mathcal{H}^{s,t} = \sum_{s<k\le t} \xi(k)$$

и

$$\mathcal{H}^{s,t} = \int_s^t \xi(u)\,du,$$

где $\xi(t)$ -некоторый случайный процесс.
Пусть

$$\eta^{s,t}(\omega) = \frac{\mathcal{H}^{s,t}(\omega) - m(s,t)}{\sigma(s,t)}$$

и $F_{\eta}^{s,t}(x)$ есть функция распределения случайной величины $\eta^{s,t}$. Нас будут интересовать условия, при которых

$$F_{\eta}^{s,t}(x) \longrightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2}\, du \qquad (*)$$

равномерно для всего семейства, когда $t-s \to \infty$.

Будем говорить, что случайные величины $\eta^{s,t}(\omega)$ удовлетворяют условию L, если для любого $\varepsilon > 0$ найдутся числа N_ε и T_ε, такие, что

$$\int\limits_{|x|>N_\varepsilon} x^2\, d F_{\eta}^{s,t}(x) \leq \varepsilon \qquad (L)$$

при $t-s \geq T_\varepsilon$.

Нетрудно понять, что условие L всегда необходимо для того, чтобы функции распределения $F_{\eta}^{s,t}(x)$ равномерно сходились к некоторому непрерывному закону с математическим ожиданием, равным 0, и единичной дисперсией. При некоторых условиях слабой зависимости и регулярности роста дисперсий $\sigma^2(s,t)$ условие L оказывается и достаточным для сходимости $F_{\eta}^{s,t}(x)$ при $t-s \to \infty$, причём к нормальному закону. Именно, имеет место следующая теорема [14].

Теорема I

Пусть величины $\mathcal{H}^{s,t}(\omega)$ являются функционалами от траекторий некоторого случайного процесса $\xi(t)$, обладающего свойством (1) сильного перемешивания, причём

$$M\left[\mathcal{H}^{s,t} - M\left(\mathcal{H}^{s,t} / \mathfrak{M}_{s-u}^{t+u}\right)\right]^2 \leq C(t-s)\,\varphi(u) \qquad (4)$$

где C -некоторая постоянная, а $\varphi(u) \to 0$ при $u \to \infty$ и кроме того[x)]

$$\sigma^2(s,t) \asymp t-s \; . \qquad (5)$$

Тогда условие L не только необходимо, но и достаточно для равномерной асимптотической нормальности (*) величин $\eta^{s,t}(\omega)$.

Качественно простое условие L к сожалению, часто бывает трудно проверить. Если же показатель $\alpha(\tau)$ в условии (1) сильного перемешивания убывает достаточно быстро,

$$\alpha(\tau) = O\left[\tau^{-1-\varepsilon}\right] \qquad (6)$$

при $\tau \to \infty$, то асимптотическая нормальность величин $\eta^{s,t}(\omega)$ вытекает из существования моментов достаточно высокого порядка [13].

---

x) Символ $\asymp$ означает, что

$$0 < \varliminf_{t-s\to\infty} \frac{\sigma^2(s,t)}{t-s} \leq \varlimsup_{t-s\to\infty} \frac{\sigma^2(s,t)}{t-s} < \infty \; .$$

Теорема 2.

Пусть величины $\mathcal{H}^{s,t}(\omega)$ измеримы относительно σ-алгебр $\mathcal{M}_s^t$ и выполняются требования (5) и (6). Если, кроме того, при некотором $\Delta > 0$

$$M|\mathcal{H}^{s,t} - m(s,t)|^{2+\delta} \leq M_0 < \infty \qquad (7)$$

где $t-s=\Delta$, а $\delta > \frac{2}{\varepsilon}$, то имеет место соотношение (*).

Отметим, что сформулированные выше результаты легко переносятся на многомерный случай / ср. [14] /

Аналогичные результаты имеют место также в случае "схемы серий", т.е. когда имеется семейство случайных функций $\mathcal{H}_n^{s,t}(\omega)$, зависящих от параметра n, $n \to \infty$, и изучается поведение функции распределения величин $\eta_n^{s,t}(\omega)$ где $s=s_n$, $t=t_n$, $t_n - s_n \to \infty$ при $n \to \infty$. / ср. ниже §5 /.

§ 2. ЛОКАЛЬНАЯ ГАУССОВОСТЬ СПЕКТРАЛЬНЫХ МЕР СТАЦИОНАРНЫХ ПРОЦЕССОВ.

В многочисленной радиотехнической литературе считалось, что после прохождения стационарного процесса через узкий линейный фильтр, он становится почти гауссовским.

Этот факт может получить строгое математическое обоснование.

Для простоты ограничимся случаем, когда частотная характеристика $\varphi_n(\lambda)$ линейного фильтра, пропускающего частоты, близкие к λ_0, имеет вид:

$$\varphi_n(\lambda) = \frac{1}{2}\left\{\varphi[(\lambda-\lambda_0)n] + \varphi[(\lambda+\lambda_0)n]\right\}\sqrt{n} \qquad (8)$$

где функция $\varphi(\lambda)$ такова, что $\varphi(-\lambda) = \overline{\varphi(\lambda)}$, $\int\limits_{-\infty}^{\infty} |\varphi(\lambda)|^2 d\lambda < \infty$ и её преобразование Фурье равномерно непрерывно ~~всюду~~ почти. Пусть действительный стационарный в широком смысле процесс $\xi(t)$,

$$\xi(t) = \int\limits_{-\infty}^{\infty} e^{i\lambda t}\, \Phi(d\lambda),$$

имеет ограниченную спектральную плотность $f(\lambda)$, такую, что

$$\operatorname{vrai}\inf_{|\lambda-\lambda_k|<\varepsilon} f(\lambda) > 0 \qquad (9)$$

где $\inf$ берется в некоторой окрестности точек λ_k, $k=\overline{1,m}$.

Рассмотрим случайные величины

$$\eta_n^{(k)} = \frac{1}{\sigma_{kn}} \int\limits_{-\infty}^{\infty} e^{i\lambda t_k}\, \varphi_n^{(k)}(\lambda)\, \Phi(d\lambda) \qquad (10)$$

где $\varphi_n^{(k)}(\lambda)$ есть частотная характеристика вида (8) с $\lambda_0 = \lambda_k$,

$$\sigma_{kn}^2 = \mathcal{D} \int\limits_{-\infty}^{\infty} e^{i\lambda t_k}\, \varphi_n^{(k)}(\lambda)\, \Phi(d\lambda).$$

Из теоремы 2 можно вывести следующий факт (ср. [11]).

Теорема 3.

Пусть стационарный процесс $\xi(t)$ обладает свойством (I) сильного перемешивания, причём

$$\alpha(\tau) = O[\tau^{-1-\varepsilon}]$$

и для некоторого $\delta > \frac{2}{\varepsilon}$

$$M|\xi(t)|^{2+\delta} \leq M_0 < \infty \qquad (11)$$

при всех t

Пусть, далее,

$$\lim_{n\to\infty} M\eta_n^{(k)}\eta_n^{(j)} = b_{kj}, \quad k,j=\overline{1,m}. \qquad (12)$$

Тогда совместная функция распределения величин $\eta_n^{(1)},\ldots,\eta_n^{(m)}$ сходится к нормальному закону с матрицей дисперсий $\|b_{kj}\|$.

Условие (I2) выполняется, например, если все частотные характеристики $\varphi_n^{(k)}(\lambda)$ отвечают одной и той же функции $\varphi(\lambda)$, фигурирующей в формуле (8), или же, например, если точки $\lambda_1,\ldots,\lambda_m$ все различны - в этом случае матрица $\|b_{kj}\|$ являются единичной; в частности, к нормальному закону с единичной матрицей моментов сходится функция распределения величин

$$\eta_n^{(k)} = \frac{2\,\mathrm{Re}\,\Phi(\Delta_n^k)}{[F(\Delta_n^k)]^{1/2}} \quad и \quad \zeta_n^{(k)} = \frac{2\,\mathrm{Im}\,\Phi(\Delta_n^k)}{[F(\Delta_n^k)]^{1/2}}$$

где $\Delta_n^k = (\lambda_k - \frac{1}{n}, \lambda_k + \frac{1}{n}), k=\overline{1,m}$; точки $\lambda_1,\ldots,\lambda_m$ все различны, и $F(\Delta) = M|\Phi(\Delta)|^2$ есть спектральная мера процесса $\xi(t)$.

§ 3. ПРИМЕНИМОСТЬ ЦЕНТРАЛЬНОЙ ПРЕДЕЛЬНОЙ ТЕОРЕМЫ К ЛОГАРИФМУ " ОТНОШЕНИЯ ПРАВОПОДОБИЯ"

Пусть вероятностная мера $P(d\omega)$, отвечающая некоторому случайному процессу $\xi(t)$, зависит от параметра, принимающего значения из некоторого интервала,

$$P(d\omega) = P_\theta(d\omega)$$

Обозначим через $P^{s,t}$ вероятностную меру, отвечающую процессу $\xi(t)$, если его рассматривать лишь на отрезке $[s,t]$ / $P^{s,t}$ определена на σ-алгебре $\mathcal{M}_s^t$ и совпадает на множествах $A \in \mathcal{M}_s^t$ с мерой P, $P^{s,t}(A) = P(A)$ /.

Пусть $m(d\omega)$ есть некоторая мера (не обязательно вероятностная), определенная на σ-алгебре $\mathcal{M}_{-\infty}^{\infty}$ такая, что $P^{s,t}$ абсолютно непрерывна относительно $m^{s,t}$ для любых s и t, и пусть

$$p^{s,t}(\omega,\theta) = \frac{P_\theta^{s,t}(d\omega)}{m^{s,t}(d\omega)}$$

есть плотность распределения вероятностей ("отношение правдоподобия").

Важным свойством является асимптотическая нормальность величин

$$L^{s,t}(\omega,\theta) = \log p^{s,t}(\omega,\theta)$$

и их производных по параметру $\frac{\partial}{\partial\theta} L^{s,t}(\omega,\theta)$ при $t-s \to \infty$.[14]

Отметим, что, как правило, можно так определить условные меры $P^{s,t}(d\omega/\mathcal{M}^s_t)$, что $P^{s,t}(d\omega/\mathcal{M}^s_{s-v})$ будет абсолютно непрерывна относительно $P^{s,t}(d\omega/\mathcal{M}^s_{s-u})$ при $u \le v$ для почти всех ω.

Положим

$$\pi^{s,t}_\theta(u,v) = \frac{P^{s,t}_\theta(d\omega/\mathcal{M}^s_{s-v})}{P^{s,t}_\theta(d\omega/\mathcal{M}^s_{s-u})}$$

и аналогично,

$$\mu^{s,t}(u,v) = \frac{m^{s,t}(d\omega/\mathcal{M}^s_{s-v})}{m^{s,t}(d\omega/\mathcal{M}^s_{s-u})}.$$

Пусть, далее,

$$\ell^{s,t}(\omega,\theta) = \frac{L^{s,t}(\omega,\theta) - ML^{s,t}}{\sqrt{DL^{s,t}}}$$

Теорема 4

Пусть свойство (I) сильного перемешивания выполняется равномерно по параметру θ,

$$DL^{s,t}(\omega,\theta) \asymp t-s \tag{13}$$

и

$$D\log\frac{\pi^{s,t}_\theta(u,v)}{\mu^{s,t}(u,v)} \le C(t-s)\varphi(u) \tag{14}$$

где C -некоторая постоянная , а $\varphi(u) \to 0$ при $u \to \infty$.

Тогда для асимптотической нормальности величин $\ell^{s,t}(\omega,\theta)$ при $t-s \to \infty$ (равномерной по параметру θ, s и t) необходимо и достаточно, чтобы они удовлетворяли условию L.

В некоторых вопросах может быть интересен случай, когда асимптотически нормальны величины $\ell^{s,t}(\omega,\theta_{s,t})$ где параметр θ меняется, когда $t-s \to \infty$. Для этого нужно лишь вместо равномерности по θ в условиях (I3) и (I4) потребовать их выполнение при $\theta = \theta_{s,t}$.

Аналогичная теорема имеет место для величин

$$\frac{\partial}{\partial\theta} L^{s,t}(\omega,\theta)$$

нужно только заменить условия (I3) и (I4) следующим:

$$\mathcal{D}\,\frac{\partial}{\partial\theta} L^{s,t}(\omega,\theta) \asymp t-s \qquad (15)$$

$$\mathcal{D}\,\frac{\partial}{\partial\theta}\log\frac{\pi_\theta^{s,t}(u,v)}{\mu^{s,t}(u,v)} \le C(t-s)\varphi(u)$$

Сформулированные выше факты вытекают из общей теоремы I, если её применить к величинам

$$\mathcal{H}_\tau^{s,t} = L^{\tau,t} - L^{\tau,s}$$

и ,соответственно, к $\frac{\partial}{\partial\theta}\mathcal{H}_\tau^{s,t}(\omega,\theta)$, зависящим от вспомогательного параметра τ.

§ 4. ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА ДЛЯ НЕКОТОРЫХ СТАЦИОНАРНЫХ ПРОЦЕССОВ.

Пусть процесс $\xi(t)$ (параметр t принимает целые значения) стационарен в узком смысле и обладает свойством (2).

Пусть случайная величина $\eta(\omega)$ измерима относительно σ-алгебры $\mathfrak{M}_{-\infty}^{\infty}$ и $\eta(t) = \eta(S_t\omega)$ есть стационарный процесс, порождаемый случайной величиной η и преобразованием сдвига S_τ^*,

$$S_\tau^* \xi(t) = \xi(t+\tau).$$

Пусть

$$M\eta = 0, \quad D\eta < \infty$$

и

$$\sum_{k=1}^{\infty} \left[M|\eta - M(\eta/\mathfrak{M}_{-k}^{k})|^2 \right]^{1/2} < \infty. \qquad (16)$$

Если функция $\beta(\tau)$ в условии (2) убывает достаточно быстро,

$$\sum_{k=1}^{\infty} \beta^{1/2}(k) < \infty \qquad (17)$$

то

$$\sigma^2 = M\eta^2 + 2\sum_{k=1}^{\infty} M\eta\cdot\eta(k) < \infty . \qquad (18)$$

Ибрагимовым [10] была установлена следующая теорема .

Теорема 5.

Если выполнены требования (I6), (I7) и (I8) , и $\sigma \neq 0$, то

$$P\left\{ \frac{1}{\sigma\sqrt{n}} \sum_{1}^{n} \eta(k) < x \right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2}\, du$$

при $n \to \infty$.

Из этого результата вытекает, в частности, центральная предельная теорема для цепей Маркова [4] (в этом случае, при широких условиях эргодичности, функция $\beta(\tau)$ убывает экспоненциально быстро).

К сожалению, свойством (2) обладают процессы, хотя и важного, но довольно узкого класса.

Так, например, гауссовские процессы входят в этот класс лишь в случае, когда корреляционная функция $B(s,t) \equiv 0$ при $|t-s| \geq T$ для некоторого $T < \infty$.

Как уже отмечалось, свойством (2) обладают эргодические марковские и m-зависимые случайные процессы.

Ибрагимовым [15] теорема 5 была применена к некоторым специальным процессам, и получен ряд новых результатов, представляющих интерес для теории чисел (см., также, [16]).

Рассмотрим произвольное число $x \in [0,1]$ и его разложение в двоичную дробь:

$$x = \frac{e_1(x)}{2} + \frac{e_2(x)}{4} + \dots + \frac{e_k(x)}{2^k} + \dots$$

Если рассматривать $e_k(x)$ как случайные величины (пространство элементарных событий Ω есть отрезок $[0,1]$, а вероятностная мера есть просто мера Лебега), то они окажутся независимыми и одинаково распределенными:

$$P\{e_k(x) = 0\} = P\{e_k(x) = 1\} = \frac{1}{2},$$

$$P\{e_1(x) = i_1, \dots, e_s(x) = i_s\} = \frac{1}{2^s}.$$

Рассмотрим стационарный процесс $\xi(t) = e_t(x),\ t \geq 0.$

Очевидно, σ- алгебра $\mathfrak{M}_0^k$ есть алгебра, натянутая на множества $A_{jk} = [j-1/2^k,\ j/2^k],$ $j = 1, \dots, 2^k.$

Пусть $f(x)$ есть произвольная функция от x такая, что

$$\int_0^1 f(x)\,dx = 0, \quad \int_0^1 f^2(x)\,dx < \infty.$$

Легко видеть, что при $x \in A_{jk}$,

$$[f]_k(x) = M\left(f/\mathfrak{M}_0^k\right) = 2^k \int_{A_{jk}} f(x)\,dx \tag{19}$$

и условие (I6) принимает вид:

$$\sum_1^\infty \left[\int_0^1 |f(x) - [f]_k(x)|^2\,dx\right]^{1/2} < \infty. \tag{20}$$

Преобразование сдвига S^* действует на Ω по формуле

$$S_t^* x = 2^t x \ (\mathrm{mod}\ 1)$$

Из теоремы I вытекает следующий результат:

<u>При условии (20)</u>

$$\sigma^2 = \int_0^1 f^2(x)\,dx + 2\sum_1^\infty \int_0^1 f(x) f(2^k x)\,dx < \infty$$

<u>и если, кроме того, $\sigma \neq 0$ то</u>

$$P\left\{\frac{1}{\sigma\sqrt{n}} \sum_1^n f(2^k x) < z\right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-u^2/2}\,du.$$

Условие же (20) будет выполнено, например, если удовлетворяется хотя бы одно из требований:

(A) $f(x)$ - функция ограниченной вариации ,

(B) $\int_0^1 |f(x) - f(x+h)|^2 dx \leq C / \log^{2+\varepsilon} h /$

где C есть некоторая постоянная, а $\varepsilon > 0$,

(C) Коэффициенты Фурье $a_n = \int_0^1 e^{2\pi n x} f(x) dx$ убывают достаточно быстро,

$$|a_n| \leq C n^{-1/2} \log^{\frac{3+\varepsilon}{2}} n .$$

Другой результат Ибрагимова относится к непрерывным дробям. Пусть попрежнему $x \in [0,1]$ и

$$x = \cfrac{1}{a_1(x) + \cfrac{1}{a_2(x) + \dots}}$$

есть его разложение в непрерывную дробь, и будем писать

$$x = [a_1(x), a_2(x), \dots].$$

Положим

$$Sx = [a_2(x), a_3(x), \dots]$$

Рассмотрим меру μ на лебеговских множествах, задаваемую формулой

$$\mu(A) = \frac{1}{\log 2} \int_A \frac{dx}{1+x} .$$

Оказывается, последовательность $\xi(t) = a_t(x)$ есть стационарный в узком смысле процесс / с вероятностной мерой μ /, обладающий свойством (2), причём соответствующая функция $\beta(\varepsilon)$ убывает очень быстро:

$$\beta(\tau) \le C e^{-\lambda\sqrt{\tau}}$$

где C -некоторая постоянная, $\lambda > 0$.

Из теоремы 5 вытекает следующий результат:

Пусть функция $f(x)$ такова, что

$$\int_0^1 f(x)\mu(dx) = 0, \quad \int_0^1 f^2(x)\,dx < \infty$$

и

$$|f(x+h) - f(x)| \le C|\log^{-1-\varepsilon} h|. \qquad (21)$$

Тогда

$$\sigma^2 = \int_0^1 f^2(x)\mu(dx) + 2\sum_1^\infty \int_0^1 f(x)f(S^k x)\mu(dx) < \infty$$

и если, кроме того, $\sigma \ne 0$, то

$$\mu\left\{x: \frac{1}{\sigma\sqrt{n}}\sum_1^n f(S^k x) < z\right\} \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-u^2/2}\,du.$$

§ 5. АСИМПТОТИЧЕСКИЕ РАЗЛОЖЕНИЯ И БОЛЬШИЕ УКЛОНЕНИЯ.

НА случай слабо зависимых величин могут быть перенесены результаты [18]-[20], уточняющие центральную предельную теорему.

Статулявичусом [17] был рассмотрен случай серий случайных величин $\xi_n(t)$, удовлетворяющих при каждом n условию (1) сильного перемешивания, в котором функции $\alpha_n(\tau)$ имеют вид:

$$d_n(\tau) = (\tau d_1)^{-a}, \quad a > 0, \; d_1 > 0. \tag{22}$$

Пусть случайные величины $\xi_n(t)$ равномерно ограничены,

$$|\xi_n(t)| \leq C \tag{23}$$

и

$$\mathcal{D} \sum_{k}^{\ell} \xi_n(t) \geq c \cdot d_n \cdot (\ell - k) \tag{24}$$

Теорема 6.

Если величины $\xi_n(t)$ удовлетворяют условиям (22), (24) и кроме того,

$$d_n \cdot n^{1/2} \to \infty \tag{25}$$

при $n \to \infty$, то при любом целом k

$$3 \leq k \leq (a-3)^{1/2}$$

для характеристической функции $f_n(u)$ случайной величины

$$\zeta_n = \frac{\sum_1^n \xi_n(t) - M \sum_1^n \xi_n(t)}{\left[\mathcal{D} \sum_1^n \xi_n(t)\right]^{1/2}}$$

в интервале $|u| \leq k\sqrt{\ln n}$ x) имеет место разложение

$$f_n(u) = e^{-u^2/2}\left(1 + \sum_{j=1}^{k-3} P_{nj}(iu)\frac{1}{\tau_n^j}\right) + O\left(\frac{|u|^k + |u|^{3k}}{\tau_n^{k-2}}\right)e^{-\frac{u^2}{2}(1+\varepsilon_n)} \tag{26}$$

x) Если $|u| > k\sqrt{\ln n}$, то

$$|f_n(u)| \leq e^{-ct^2}$$

при некотором $c > 0$.

где

$$\tau_n = \frac{d_n^2}{n}\left[\mathcal{D}\sum_1^n \xi_n(t)\right]^{3/2},$$

$$c_1 d_n \cdot n \leq \mathcal{D}\sum_1^n \xi_n(t) \leq c_2 \cdot n \cdot \frac{1}{d_n}$$

а коэффициенты полиномов $P_{nj}(it)$ равномерно ограничены по n, компоненты в символе " O " зависят от k и ограниченных равномерно по n, $\varepsilon_n \to 0$ при $n \to \infty$.

Далее, если показатель α_n в условии (22) равномерно больше нуля,

$$\alpha_n \geq \alpha > 0 \qquad (27)$$

и кроме того, постоянная $a \geq 12$, то для вероятности

$$P\left\{\sum_1^n \xi_n(t) > x\right\} = 1 - F_n(x)$$

больших уклонений $1 \leq x \leq o\sqrt{n}$ имеет место следующее соотношение Крамера-Петрова:

$$\frac{1 - F_n(x)}{\int_x^\infty e^{-u^2/2}\,du} = e^{x^3/\sqrt{n}} \cdot \lambda_n(x/\sqrt{n})\left(1 + O(x/\sqrt{n})\right) \qquad (28)$$

где степенной ряд $\lambda_n(y)$ сходится при малых y равномерно по n.

ЛИТЕРАТУРА.

1. С.Н.Бернштейн, Распространение предельной теоремы теории вероятностей на суммы зависимых величин, УМН, X, 65-115, 1944 /впервые опубликована в Math. Annalen, 97, 1-59, 1926/.

2. С.Х.Сираждинов, Предельные теоремы для однородных цепей Маркова, Ташкент, 1955.

3. Р.Л.Добрушин, Центральная предельная теорема для неоднородных цепей Маркова, Теор. вер. и ее прим., I.№ I, 72-89, 1956.

4. С.В.Нагаев, Некоторые предельные теоремы для однородных цепей Маркова, Теор.вер. и ее прим., 2 № 4, 389-416, 1957

5. В.А.Статулявичус, Асимптотические разложения для неоднородных цепей Маркова, ДАН СССР, 112,№2, 1957

6. W. Höffding, H. Robbins, The central limit theorem for dependent random variables, Duke Math. J., 15, 773-780, 1948

7. P. H. Diananda, The central limit theorem for m-dependent variables asymptotically stationary to second order, Proc. Cambr. Phil. Soc., 50, 287-292, 1954.

8. Каллианпур (G. Kallianpur), Об одной предельной теореме для зависимых случайных величин, ДАН СССР, IOI, I3-I6, I955

9. M. Rosenblatt, A central limit theorem and a strong mixing condition, Proc. Nat. Acad. Sci. Wash., 42, n 3, 43-47, 1956.

IO. И.А.Ибрагимов, Некоторые предельные теоремы для стационарных в узком смысле вероятностных процессов, ДАН СССР, I25, 4, I959.

II. В.А.Волконский, Ю. А.Розанов, Некоторые предельные теоремы для случайных функций, I, Теор.вер. и ее прим., 4, №2 I86-207, I959.

I2. А.Н.Колмогоров, Ю.А.Розанов, О свойстве сильного перемешивания гауссовского стационарного процесса, Теор. вер., и ее прим., 5, № 2, I960

I3. Ю.А. Розанов, О центральной предельной теореме для аддитивных случайных функций, Теор. вер.и ее прим., 5, № 2, I960

I4. Ю.А.Розанов, О центральной предельной теореме и ее применимости к отношению правдоподобия (принята к опубликованию в "Теор.вер. и ее прим.")

I5. И.А.Ибрагимов, Асимптотическое распределение некоторых сумм, Весник ЛГУ, № I, вып.I, I960, 55-69

I6. M. Kac, On distribution of values of sums of the type $\sum_{0}^{n-1} f(2^k t)$, Ann. Math., N 47, 33-49, 1946.

I7. В.А.Статулявичус, Доклад на Всесоюзной конференции по теории вероятностей, г. Ужгород, сентябрь 1959; резюме доклада опубликовано в Теор. вер. и ее прим., 5, N 2, 1960.

I8. Б.В.Гнеденко, А.Н.Колмогоров, Предельные распределения для сумм независимых случайных величин, М.-Л. 1949

I9. H. Cramer, Sur un nouveau théorème-limite de la théorie des probabilités, Actualités scientifiques et industrielles, Paris, 1938.

20. В.В.Петров, Распространение теоремы Крамера на неодинаково распределенные независимые величины, Вести. ЛГУ, № 8, 1958.

CHARACTERIZATION OF SAMPLE FUNCTIONS OF STOCHASTIC PROCESSES BY SOME ABSOLUTE PROBABILITIES

Marek Fisz

University of Warsaw

Mathematical Institute

Polish Academy of Sciences

1. Introduction

The investigation of properties of sample functions of different stochastic processes has attracted much attention. It seems, indeed, that the characterization of a stochastic process is not exhaustive unless such basic properties of sample functions as continuity, kind of discontinuity (if any), integrability and so on, are known. The most advanced investigations in this field are for certain particularly distinguished classes of stochastic processes, namely Markov processes, processes with independent increments and martingales. There are important results due to Doeblin, Doob, Levy, Wiener, and others. For arbitrary stochastic processes, without the assumption that they belong to some traditionally distinguished class of stochastic processes, conditions have been given, expressed in terms of the moments of the random variables of the processes considered, under which almost all sample functions are continuous (Kolmogorov, [11]) or have no discontinuities of the second kind [12]. The author [7] has given conditions expressed in terms of some absolute probabilities under which almost all sample functions of the process are step functions with a finite expected number of discontinuities. The object of this note is to strengthen these results and to obtain other related results.

2. Notation and summary

We consider the real separable (see p. 51, [4]) stochastic process $\{x_t, t \in I_0\}$ where I_0 is a closed interval. We denote by Ω the set of elementary events ω, by $\mathcal{B}$ the smallest Borel field of ω-sets with respect to which all the random variables x_t where $t \in I_0$, are measurable, and by P the probability measure on $\mathcal{B}$. As is known [9] the Borel field $\mathcal{B}$ is generated by the aggregate of ω-sets of the form $\{[x_{t_1}(\omega), \ldots, x_{t_n}(\omega)] \in A\}$, where A is any right semiclosed interval, and $(t_1, \ldots, t_n)$ is any finite set of values of $t \in I_0$ and $n = 1, 2, \ldots$. The measure P on $\mathcal{B}$ is uniquely determined by the finite-dimensional probability measures of ω-sets of the above form. Denote by x_I the increment of x_t in the interval $I \subset I_0$ and by $|I|$ the length of I. We shall consider the following functions.

$$a(I) = P\{x_I \neq 0\}, \tag{1}$$

$$b(I, \varepsilon) = P\{|x_I| > \varepsilon\}, \tag{2}$$

(3) $$A(I) = \int_I a(J) = \lim_{n\to\infty} \sum_{k=1}^{n} a(I_{nk})$$

(4) $$B(I, \varepsilon) = \int_I b(J,\varepsilon) = \lim_{n\to\infty} \sum_{k=1}^{n} b(I_{nk}\ \varepsilon)$$

as $\max\limits_{1\le k\le n} |I_{nk}| \longrightarrow 0$, where $\{I_{nk}\}$ is a partition of I into nonoverlapping intervals I_{nk} $(k = 1, 2,\ldots, n)$. The expressions $A(I)$ and $B(I,\varepsilon)$ are the Burkill integrals of $a(I)$ and $b(I,\varepsilon)$ respectively. The corresponding upper Burkill integrals $\bar{A}(I)$ and $\bar{B}(I,\varepsilon)$ are obtained by replacing in (3) and (4) the symbol lim by $\overline{\lim}$. Finally we define

(5) $$Q(t) = \lim \frac{a(I)}{|I|}$$

as I contracts to a fixed point $t \in I$.

Conditions expressed in terms of the functions (1) to (5) are given under which almost all sample functions of the process

(a) are jump functions where locally monotonic, as defined below,

(b) have no discontinuity of the first kind at a given but arbitrary point $t \in I_0$,

(c) are step functions without fixed discontinuities and with a finite expected number of discontinuities equal to $\bar{A}(I)$,

(d) are constant on the whole interval I_0,

(e) have no discontinuities of the first kind.

A condition is given in terms of the covariance function of a process stationary in the wide sense, under which almost all sample functions are continuous.

3. General results

We begin with the following definitions.

Definition 1. The sample function $x_t(\omega)$ will be called a jump function where locally monotonic (briefly, a jump function wlm) if at any continuity point $t \in I_0 - S(\omega)$ of $x_t(\omega)$, where $S(\omega)$ is some possibly empty set nowhere dense in I_0, at which $x_t(\omega)$ is locally monotonic, the equality $x_{t'}(\omega) = x_t(\omega)$ holds for all t' of some interval $(t - \theta(t,\omega), t + \theta(t,\omega))$.

Definition 2. The sample function $x_t(\omega)$ will be called a jump function if at any continuity point $t \in I_0$ of $x_t(\omega)$ the equality $x_{t'}(\omega) = x_t(\omega)$ holds for all t' from some interval $(t - \theta(t,\omega), t + \theta(t,\omega))$.

It is evident how definitions 1 and 2 should be modified for the endpoints of the interval I_0.

Theorem 1. Let the process $\{x_t, t \in I_0\}$ be seperable and let the relation

(6) $$\lim_{|I| \to 0} a(I) = 0$$

hold where $I \subset I_0$. Then almost all (P) sample functions of the process are jump functions wlm.

Proof. Let relation (6) hold. We can then find for any fixed point $t \in I_0$ a sequence of parameter points $t'_r \neq t$ for $r = 1, 2, \ldots,$ converging to t and such that the relations

$$(7) \qquad P\{x_t - x_{t'_r} \neq 0\} < \frac{1}{r^2}$$

hold for $r = 1, 2, \ldots$. In virtue of the Borel-Cantelli lemma the ω-set Λ_t which is such that for any $\omega \in \Lambda_t$ there exists an $R(\omega)$ such that for all $r > R(\omega)$ the equality $x_t(\omega) = x_{t'_r}(\omega)$ holds, satisfies the equality

$$(8) \qquad P(\Lambda_t) = 1.$$

Denote by Δ_t the ω-set such that for any $\omega \in \Delta_t$ there exists for every $\theta > 0$ a point $t'(\omega)$ of the interval $(t - \theta, t + \theta)$ satisfying the equality $x_t(\omega) = x_{t'}(\omega)$. Since $\Lambda_t \subset \Delta_t$ we have by virtue of (8) the equality

$$(9) \qquad P(\Delta_t) = 1.$$

Let now $T = \{t_j\} (j = 1, 2, \ldots)$ be some denumerable dense set in I_0. Relation (9) implies

$$(10) \qquad P\left(\bigcap_{j=1}^{\infty} \Delta_{t_j}\right) = 1.$$

Since the process is separable and, in virtue of (6), T satisfies the separability conditions it follows from (10) that almost all sample functions $x_t(\omega)$ are jump functions wlm.

Before formulating theorem 2 let us remember that the function $x_t(\omega)$ is said to have a discontinuity of the first kind at some point t if $x_t(\omega)$ has at t both onesided limits and they are not equal. If at least one onesided limit at t does not exist the discontinuity is said to be of the second kind.

Theorem 2. Let $\{x_t, t \in I_0\}$ be a real, separable stochastic process and let the relation

$$\lim_{|I| \to 0} b(I, \varepsilon) = 0 \tag{11}$$

hold for any $\varepsilon > 0$, where $I \subset I_0$. Then almost all sample functions have no discontinuity of the first kind at any fixed but arbitrary point $t \in I_0$.

Proof. Let relation (11) hold and let the assertion of the theorem not be true. We have then for some point $t \in I_0$ (different from both endpoints of I_0) the equality

$$P\{\lim_{t \uparrow t_0} x_t = x_{t_0 - 0} \neq x_{t_0 + 0} = \lim_{t \downarrow t_0} x_t\} = \alpha \tag{12}$$

where $\alpha > 0$.

Denote by $J \subset I_0$ an interval having t_0 as its midpoint, by $\Lambda_n(J)$ an ω-set for which $|x_J(\omega)| > 1/n$, and by $\Lambda(J)$ on ω-set for which $x_J(\omega) \neq 0$. The sequence $\{\Lambda_n(J)\}$ is increasing, hence

$$\begin{gathered} \Lambda(J) = \bigcup_{n=1}^{\infty} \Lambda_n(J) = \lim_{n \to \infty} \Lambda_n(J), \\ P\{x_J \neq 0\} = P\{\Lambda(J)\} = \lim_{n \to \infty} P\{\Lambda_n(J)\}. \end{gathered} \tag{13}$$

Taking into account (12) and (13) we obtain

$$\alpha \leq \lim_{|J| \to 0} P\{x_J \neq 0\} = \lim_{|J| \to 0} \lim_{n \to \infty} P\{\Lambda_n(J)\}.$$

Consequently for $|J|$ sufficiently small and n sufficiently large, that is, for ε sufficiently small, $P\{\Lambda_n(J)\}$ will be at least near α, which contradicts relation (11). Hence $\alpha = 0$, and theorem 2 is thus proved.

Let us remark that if (11) holds, any set of parameter points dense in I_0 satisfies the separability conditions, and consequently for any fixed but arbitrary $t \in I_0$ the equation

$$P\{x_{t-0} = x_{t+0} \neq x_t\} = 0 \tag{14}$$

holds. Now let (11) be satisfied and let the process $\{x_t, t \in I_0\}$ have a fixed discontinuity point t', that is, for some $\beta > 0$, the equation

$$P(\lim_{t \to t'} x_t = x_{t'}) = 1 - \beta \tag{15}$$

holds. We then obtain from theorem 2 and (14) the following.

Corollary. Let the real, separable process $\{x_t, t \in I_0\}$ satisfy (11) and let it have a fixed discontinuity point $t' \in I_0$. Then the probability that x_t has at t' a discontinuity of the second kind equals β where β is given by relation (15).

Theorem 3 which we shall now formulate is a stronger version of some results obtained by the author [7]. It was assumed in that paper that the process under consideration had no fixed discontinuities. Now this assumption is entirely omitted in part (b) and is replaced in part (a) of theorem 3 by the weaker assumption (11). The former assumption is now obtained as a consequence (ii). In obtaining these strengthened results essential use will be made of the corollary to theorem 2.

Theorem 3. Let the stochastic process $\{x_t, t \in I_0\}$ be separable.

(a) If relation (11) holds for any $\varepsilon > 0$ and if moreover

(16) $$\bar{\lambda}(I_0) < \infty$$

then for every open interval $I \subset I_0$.

(i) almost all (P) sample functions of the process are jump functions,

(ii) the process has no fixed points of discontinuity,

(iii) the relation

(17) $$E\xi(I) = \bar{\lambda}(I) = A(I)$$

holds, where $\xi(I)$ is the number of discontinuities of $x_t(\omega)$ on the interval I,

(iv) the ω-set such that at the discontinuities, if any of $x_t(\omega)$ both onesided limits exist, has probability 1,

(v) the function Q(t) exists almost everywhere in I and satisfies the relation

(18) $$\int_I Q(t)dt \leq A(I).$$

(b) If $a(I)$ is an absolutely continuous function of an interval or $a(I)$ satisfies the Lipschitz condition, then the assertions (i) to (v) are satisfied, and moreover

(19) $$\int_I Q(t)dt = A(I).$$

Proof. In the proof of assertion (i) given formerly in [7] the assumption that the process has no fixed points of discontinuity was used only for stating that any denumerable dense set of points $t \in I$ satisfies the separability conditions. Now relation (11) is sufficient for this purpose and therefore assertion (i) is proved. Moreover, in proving assertion (i), in [7] it was also proved that almost every sample function $x_t(\omega)$ has a finite number of discontinuities. Hence by virtue of relation (11) and the corollary to theorem 2, the process has no fixed points of discontinuity, which proves assertion (ii). Once this fact has been established the proof of the remaining assertions of part (a) given in [7] remains valid.

In order to prove part (b) of the theorem we remark that if $a(I)$ satisfies the Lipschitz condition it is (see p. 287, [2]) an absolutely continuous function of an interval and this implies (see p. 287, [2]) that relation (16) holds. Since (11) evidently holds then also, all the assumptions of part (b) are fulfilled and consequently so are all its assertions. Moreover the continuity of $a(I)$ implies (see p. 289, [2]) the absolute continuity of $A(I)$, hence relation (19) is true. Theorem 3 is thus proved.

Theorem 4. Let $\{x_t, t \in I_0\}$ be a real, separable stochastic process. The relation

$$P\{x_t = \text{const.}, t \in I_0\} = 1 \tag{20}$$

holds if and only if the relation

$$Q(t) \equiv 0 \tag{21}$$

holds uniformly with respect to $t \in I_0$.

Proof. Let relation (21) hold uniformly with regard to $t \in I_0$. We shall show that $a(I)$ is then an absolutely continuous function of an interval. Indeed, relation (21) implies that for any $\alpha > 0$ there exists such a $\beta > 0$ that if $|I| < \beta$ the inequality $a(I) < \alpha|I|$ is true. Let us now divide the interval I into a finite number, say n, of non-overlapping intervals I_k with $|I_k| < \beta$ for $k = 1, 2, \ldots, n$. The function $a(I)$ then satisfies the Lipschitz condition in every interval I_k and consequently $a(I)$ is an absolutely continuous function of an interval in I_k. It then follows that $a(I)$ has the same property in the whole interval I_0. By virtue of part (b) of theorem 3 the relation (19) holds. Taking into account (21) we obtain $A(I) = 0$, which in virtue of (17) implies relation (20). The "if" assertion is thus proved, the "only if" assertion is evident.

Let us remark that if we assume that for any $\varepsilon > 0$ relation (21) holds and moreover $\bar{h}(I,\varepsilon) < \infty$, we cannot obtain any new properties of the sample functions of the process in addition to the property established in theorem 2. However, the following theorem is true.

<u>Theorem 5.</u> <u>Let the stochastic process</u> $(x_t, t \in I_0)$ <u>be separable and let the relation</u>

$$\bar{h}(I,\varepsilon) = 0 \tag{22}$$

<u>hold for any</u> $\varepsilon > 0$. <u>Then almost all</u> (P) <u>sample functions</u> $x_t(\omega)$ <u>have no discontinuities of the first kind</u>.

<u>Proof</u>. Let us at first remark that relation (21) follows from relation (22). Let $T = \{t_j \mid j = 1, 2, \ldots\}$ be a denumerable dense set in I_0 and let $\tau_{n1}, \ldots, \tau_{nn}$ for $n = 1, 2, \ldots$ denote the points $t_1, \ldots, t_n$ arranged in increasing order, $\tau_{n1} < \ldots < \tau_{nn}$. Denote by I_{nk} the interval $[\tau_{nk}, \tau_{n,k+1})$ and by $\rho_n(T,\varepsilon)$ the number of intervals I_{nk} for which $|x_{I_{nk}}| > \varepsilon$. We have

$$E\rho_n(T,\varepsilon) = \sum_{k=1}^{n} h(I_{nk},\varepsilon). \tag{23}$$

It follows from relation (22) that as $\max_{1 \le k \le n} |I_{nk}| \longrightarrow 0$ we have $\lim_{n \to \infty} E\rho_n(T,\varepsilon) = 0$. The last relation implies the existence of a subsequence, $n_j \uparrow \infty$, for which the relation

$$P\{\lim_{j \to \infty} \rho_{n_j}(T,\varepsilon) = 0\} = 1 \tag{24}$$

is true. Since ρ_n can take only noninteger values, relation (24) implies that the ω-set $A(\varepsilon)$, which is such that for any $\omega \in A(\varepsilon)$ there exists a j_0 which may depend on ε and ω such that for $j > j_0$ the equality $\rho_{n_j}(T,\varepsilon) = 0$ holds, satisfies the relation $P\{A(\varepsilon)\} = 1$. Hence, taking into account the method of constructing the intervals I_{nk}, we see that almost all (P) $x_t(\omega)$ have no discontinuities of the first kind with a jump greater than ε. Consider now a sequence of positive constants $\{\varepsilon_n\}$ where $\varepsilon_n \downarrow 0$. The set Λ, such that the $x_t(\omega)$ corresponding to $\omega \in \Lambda$ have no discontinuities of the first kind, is given by the formula

$$(25) \qquad \Lambda = \bigcap_{n=1}^{\infty} A(\varepsilon_n).$$

Consequently, since $P\{A(\varepsilon_n)\} = 1$, we obtain $P(\Lambda) = 1$. Theorem 5 is thus proved.

Let us remark that, as Dobrushin [3] has proved, the assertion of theorem 5 holds under the assumption that as $\Delta t \longrightarrow 0$ the relation

$$(26) \qquad \sup_{\substack{t \in I_0 \\ t+\Delta t \in I_0}} P\{|x_{t+\Delta t} - x_t| > \varepsilon\} = o(\Delta t)$$

holds for any $\varepsilon > 0$. We shall show that relation (26) implies relation (22). Indeed it follows from (26) that for any $\eta > 0$ there exists such an $\alpha > 0$ that for any

$t \in I_0$, $t + \Delta t \in I_0$ with $|\Delta t| \leq \alpha$ we have $P(|x_{t+\Delta t} - x_t| > \varepsilon) < \eta|\Delta t|$. In other words for sufficiently large n we shall have for $k = 1, \ldots, n$ the relation $b(I_{nk}, \varepsilon) < \eta|I_{nk}|$ and consequently

(27) $$\sum_{k=1}^{n} b(I_{nk}, \varepsilon) < \eta|I_0|.$$

Since η is arbitrary small we obtain relation (22).

It is easy, however, to show that relation (26) is a more restrictive condition than (22). Indeed the following example shows that (22) may be satisfied although (26) is not.

Example. Let $\Omega = \{\omega_j | j = 1, 2, \ldots\}$ and let $P(\omega_j) = 1/2^j$. Let $x_t(\omega_j)$ be given by the formula

(28) $$x_t(\omega_j) = \begin{cases} 2^j t & 0 \leq t \leq 1/2^j \\ 1 & 1/2^j < t \leq 1 \end{cases} .$$

It is evident that for this process relation (26) does not hold. We shall show that relation (22) holds. Let $\eta > 0$ be an arbitrary number. We have to show that for sufficiently small $\max\limits_{1 \leq k \leq n} |I_{nk}|$ the inequality

(29) $$\sum_{k=1}^{n} P\{|x_{I_{nk}}| > \varepsilon\} < \eta$$

will be true. We have

(30) $$\sum_{k=1}^{n} P\{|x_{I_{nk}}| > \varepsilon\} = \sum_{j=1}^{\infty} \frac{1}{2^j} \sum_{k=1}^{n} P\{|x_{I_{nk}}| > \varepsilon | x_t = x_t(\omega_j)\}.$$

Let j_0 be an integer such that $(1/2^{j_0}) < \varepsilon\eta$ and consider the all parenthesis necessary? Jan 3

partition $\{I_{nk}\}$ of the interval $[0,1]$ such that $\max_{1 \le k \le n} |I_{nk}| < (\varepsilon|2^{j_0}\theta)$. We have then for $j = 1, 2, \ldots, j_0$

$$(31) \qquad P(|x_{I_{nk}}| > \varepsilon \,|\, x_t = x_t(\omega_j)) = 0.$$

Consequently

$$(32) \qquad \sum_{k=1}^{n} P(|x_{I_{nk}}| > \varepsilon) = \sum_{j=j_0+1}^{\infty} \frac{1}{2^j} \sum_{k=1}^{n} P(|x_{I_{nk}}| > \varepsilon | x_t = x_t(\omega_j)).$$

Since, as follows from the definition of the $x_t(\omega_j)$, the second sum, where k runs from 1 to n, on the right side of (32) can only be at most equal to $1/\varepsilon$, and taking into account the method of choosing j_0, relation (32) implies relation (29). Hence relation (22) holds.

4. Particular cases

We shall deal in this section with stochastic processes of certain special types.

Theorem 6. Let $\{x_t, t \in I_0\}$ be a real, separable martingale and let relation (22) hold. Then almost all (P) sample functions $x_t(\omega)$ of the process are continuous.

Proof. The assertion of the theorem follows from theorem 5 and from a theorem of Doob (see chap. 7 [4]) stating that the set of sample functions $x_t(\omega)$ of a separable martingale which have only discontinuities of the first kind, if any, has probability 1.

Dobrushin [3] has proved the assertion of theorem 6 under the stronger assumption (26).

We shall now strengthen slightly a theorem of Dynkin [5] and Kinney [9]. Denote

$$(33)\qquad C(h,\varepsilon) = \sup_{\substack{t\in I_0\\ t+\Delta t\in I_0\\ |\Delta t|\leq h\\ -\infty<x<\infty}} P\{|x_{t+\Delta t} - x_t| > \varepsilon | x_t = x\},$$

$$(34)\qquad D(h,\varepsilon) = \sup_{\substack{t\in I_0\\ t+\Delta t\in I_0\\ |\Delta t|\leq h}} \int_{-\infty}^{\infty} P\{|x_{t+\Delta t} - x_t| > \varepsilon | x_t = x\} dF_t(x),$$

where $F_t(x) = P(x_t = x)$.

<u>Theorem</u> 7. <u>Let</u> $\{x_t,\ t \in I_0\}$ <u>be a separable Markov process</u>. <u>If as</u> $h \to 0$ <u>the relation</u>

$$(35)\qquad D(h,\varepsilon) = o(h)$$

<u>holds, then almost all sample functions of the process are continuous</u>.

<u>Proof</u>. We have

$$(36)\qquad P\{|x_{t+\Delta t} - x_t| > \varepsilon\} = \int_{-\infty}^{\infty} P\{|x_{t+\Delta t} - x_t| > \varepsilon | x_t = x\} dF_t(x).$$

Relations (35) and (36) imply relation (26), hence, <u>a fortiori</u>, relation (22). By a theorem of Fuchs [8] the assertion of theorem 7 is obtained.

We remark that Dynkin and Kinney have proved the assertion of theorem 7 under the more restrictive assumption that $C(h,\varepsilon) = o(h)$.

Theorem 8 below is known. However, we shall give a simple proof using tools applied in this paper.

Theorem 8. Let $\{x_t, t \in I_0\}$ be a separable stochastic process with independent increments and let relation (6) hold. Then the assertions (i) to (v) of theorem 3 are true.

Proof. It has been proved in [6] that the independence of increments and relation (6) imply the relation (16). Consequently if the assumptions of theorem 8 are satisfied, all the assumptions of part (a) of theorem 3 are satisfied and hence all its assertions are true.

Theorem 9. Let $\{x_t, -\infty < t < +\infty\}$ be a real, separable stochastic process, stationary in the wide sense, and let the covariance function $R(\tau)$ be for $\tau \to 0$ of the form

$$R(\tau) = D^2(x_0) + 0(|\tau|^{1+\delta}) \tag{37}$$

where $\delta > 0$. Then almost all sample functions of the process are continuous.

Proof. We have for any real t and τ the relation

$$E(x_{t+\tau} - x_t)^2 = 2[D^2(x_0) - R(\tau)]. \tag{38}$$

Taking into account (37) we obtain for arbitrary t

$$E(x_{t+\tau} - x_t)^2 = 0(|\tau|^{1+\delta}). \tag{39}$$

By the result of Kolmogorov (published in [11]) relation (39) implies that almost all sample functions $x_t(\omega)$ of the process are continuous in every finite closed interval. Theorem 9 is thus proved.

We remark that if the covariance function $R(\tau)$ is twice differentiable at $\tau = 0$, then almost all sample functions of the process are absolutely continuous (see p. 536, [4]).

We remark finally that--as Belayev [1] has shown--if the process $\{x_t, -\infty < t < \infty\}$ is real, separable, stationary and Gaussian and if the relation

$$(40) \qquad R(\tau) = D^2(x_0) + O\left(\frac{1}{l_n|\tau|^{1+\delta}}\right)$$

with $\delta > 0$ holds, then almost all sample functions are continuous. Relation (40) is evidently less restrictive than relation (37), however Belayev considers only Gaussian stationary processes.

REFERENCES

[1] V.R. BELAYEV, "Continuity and Holder's condition for sample functions of stationary Gaussian processes," Proc. 4th Berkeley Symp. on Prob. and Math. Stat., Univ. Calif. Press, 1961.

[2] J.C. BURKILL, "Functions of intervals," Proc. London Math. Soc., Vol. 22 (1924), pp. 275-310.

[3] R.L. DOBRUSHIN, "The continuity condition for sample martingale functions," (in Russian), Teor. Veroyatnost. i Primenen., Vol. 3 (1958) pp. 97-98.

[4] J.L. DOOB, Stochastic Processes, New York, John Wiley and Sons, Inc., 1953.

[5] E.B. DYNKIN, "Conditions of continuity and lack of discontinuities of the second kind for the realizations of a Markov stochastic process" (in Russian), Izv. Akad. Nauk. SSSR Ser. Mat., Vol. 16 (1952), pp. 563-572.

[6] M. FISZ and K. URBANIK, "Analytic characterization of a composed nonhomogeneous Poisson process," Studia Math., Vol. 15 (1956), pp. 328-336.

[7] M. FISZ, "Realizations of some stochastic processes," Studia Math., Vol. 15 (1956), pp. 359-364.

[8] A. FUCHS, "Sur la continuité stochastique des processus stochastiques réels de Markoff," C.R. Acad. Sci. Paris, Vol. 237 (1953), pp. 1396-1399.

[9] J.R. KINNEY, "Continuity properties of sample functions of Markov processes," Trans. Amer. Math. Soc., Vol. 74 (1953), pp. 280-302.

[10] A. KOLMOGOROFF, Grundbegriffe der Wahrscheinlichkeitsrechnung, Berlin, Springer, 1933.

[11] E. SLUTSKY, "Alcuni proposizioni sulla theoria degli funzioni aleatorie," Giorn. Ist. Ital. Attuari, Vol. 8 (1937), pp. 183-199.

[12] N.N. Tschentzov, "Weak convergence of stochastic processes whose sample functions have no discontinuities of the second kind and the so-called 'heuristic approach' to the nonparametric tests of the Kolmogorov-Smirnov type" (in Russian), Teor. Veroyatnost. i Primenen., Vol. 1 (1956), pp. 154-161.

REMARKS ON PROCESSES OF CALLS

By

Czesław Ryll-Nardzewski

University of Wroclaw,

Mathematical Institute,

Polish Academy of Sciences

1. Introduction

The theory of processes of calls is highly developed. In this paper I am going to consider some questions which, to my mind, have not yet been analyzed sufficiently from the measure theoretic point of view.

Palm [1] dealt with a special kind of conditional probability for stationary processes. A. J. Khinchine [2] presented and completed the ideas of Palm. Their methods were simple and elegant but they were of analytical character. In this paper I am going to give a different, and so far as I know, new approach to these problems. I shall confine myself to considering some basic notions and their properties.

As a byproduct I have obtained a result from ergodic theory which seems to be of some interest in itself.

2. Discrete time

From the measure theoretic point of view, the theory of stationary processes of calls with discrete time is quite simple and consequently it is not dangerous to omit some technical details. Let us consider a doubly infinite stationary sequence of random variables $\cdots, \xi_{-2}, \xi_{-1}, \xi_0, \xi_1, \xi_2, \cdots$ taking only the value zero or one. It is easy to prove that either there are no calls or they occur infinitely many times in both time directions. In symbols

(1) $$P\{\xi_i \equiv 0 \quad \text{or} \quad \overline{\lim_{i\to+\infty}}\, \xi_i = \overline{\lim_{i\to-\infty}}\, \xi_i = 1\} = 1,$$

where P denotes the probability measure. The first possibility is uninteresting from any point of view. Hence we may suppose that

(2) $$P\{\xi_i \equiv 0\} = 0.$$

Further, the general case can be reduced to this case by the introduction of a "new" probability measure, invariant under the shift transformation,

(3) $$P^*(\cdot) \overset{\mathrm{df}}{=} P(\cdot|N), \quad \text{where} \quad N = \{\xi_i \equiv 0\}.$$

We denote by A the event $\{\xi_0 = 1\}$ and put $\alpha = P(N)$. Under the condition that A has occurred, that is, that there is a call at time $t = 0$, we can define the sequence of random variables $\cdots, \eta_{-2}, \eta_{-1}, \eta_0, \eta_1, \cdots,$ which are equal to the distances of the successive calls. The enumeration begins from the call ξ_0. This is illustrated in Figure 1.

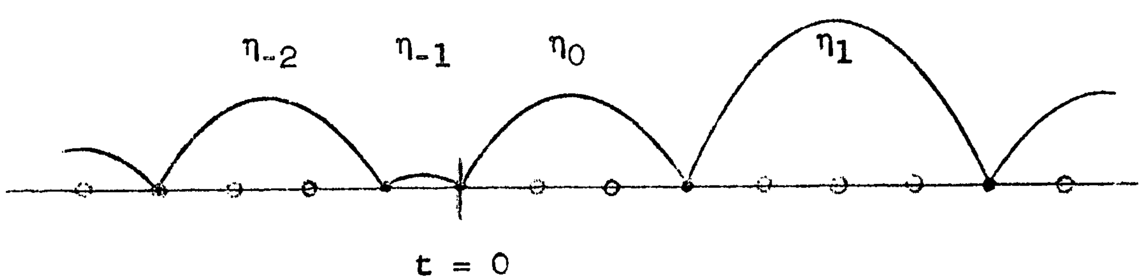

Figure 1

We denote by $p(\cdot)$ the conditional probability $p(\cdot) = P(\cdot|\xi_0 = 1) = P(\cdot|A)$. We can now assert

Theorem 1. (i) The sequence $\{\eta_i\}$ is

(a) stationary with respect to measure p,

(b) the random variables η_i admit only positive, integer values,

(c) the expectation of η_i is finite (and is equal to $(P\{\xi_0 = 1\})^{-1}$).

(ii) The correspondence between $\{\xi_i\}$ and $\{\eta_i\}$, or more precisely between the probability measures P and p, is 1 to 1.

(iii) Each sequence $\{\eta_i\}$ of random variables satisfying (a) to (c) can be obtained in this way.

Proof. Let $S_{-m}, S_{-m+1}, \cdots, S_0, S_1, \cdots, S_n$ be arbitrary positive numbers. In order to prove the equality

$$(4) \qquad P\{\eta_{-m} = S_{-m}, \cdots, \eta_n = S_m\} = p\{\eta_{-m+1} = S_{-m}, \cdots, \eta_{n+1} = S_n\}$$

it is sufficient to apply the shift transformation S_{-1} times in the formula

$$(5) \qquad P\{\xi_{s_{-m}} = 1, \xi_{s_{-m}+1} = 0, \ldots, \xi_{s_{-m+1}} = 1, \ldots, \xi_{s_{-1}} = 1, \xi_{s_{-1}+1} = 0, \ldots, \xi_{-1} = 0,$$
$$\xi_0 = 1,\ \xi_1 = 0, \xi_{s_1} = 1, \ldots, \xi_{s_n} = 1\}.$$

Hence property (a) has been proved. Property (b) is obvious. Further, by the stationarity of $\{\xi_n\}$ and formulas (1) and (2) we have

$$(6) \qquad E_p(\eta_0) = \sum_{i=1}^{\infty} p\{\eta_0 \geq i\} = \frac{1}{\alpha}\sum_{i=1}^{\infty} P\{\xi_0 = 1, \xi_1=0,\cdots,\xi_{i-1}=0\}$$

$$= \frac{1}{\alpha}\sum_{i=0}^{\infty} P\{\xi_{-i}=1,\xi_{-i+1}=0,\cdots,\xi_{-1}=0, \xi_0=0\} = \frac{1}{\alpha}.$$

Thus (e) is proved, which completes the proof of the first part of our assertion.

Now we shall consider the correspondence between the probability measures P and p. We have

$$(7) \qquad P(\mathcal{E}) = \sum_{i=0}^{\infty} P\{\xi_{-i}=1,\xi_{-i+1}=0,\cdots,\xi_0=0, \mathcal{E}\}$$

$$= \sum_{i=0}^{\infty} P\{\xi_0=1,\xi_1=0,\cdots,\xi_i=0, \mathcal{E}^{(i)}\},$$

where $\mathcal{E}^{(i)}$ denotes the event $\mathcal{E}$, shifted to the right i times. Finally, we obtain from (7) the following formula expressing the probability measure P by the probability measure p

$$(8) \qquad P(\mathcal{E}) = \alpha \sum_{i=0}^{\infty} p\{\eta_0 > i, \mathcal{E}^{(i)}\}.$$

Now we have to prove only the last part of theorem 1. Namely, we suppose that conditions (a) to (c) hold, and formula (8) can be regarded as a definition of the measure P. Evidently we must put $\alpha \overset{df}{=} (\int \eta_0 \, dp)^{-1}$. We shall compute the value of $P(\mathcal{E}^{(1)})$. We have

$$(9) \qquad P(\mathcal{E}^{(1)}) = \alpha \sum_{i=0}^{\infty} p\{\eta_0 > i, \mathcal{E}^{(i+1)}\}$$

$$= \alpha \sum_{i=0}^{\infty} \left[p\{\eta_0 > i, \mathcal{E}^{(i)}\} + p\{\eta_0=i, \mathcal{E}^{(i)}\}\right] - \alpha p(\mathcal{E})$$

$$= P(\mathcal{E}) + \alpha\left[\sum_{i=0}^{\infty} p\{\eta_0=i, \mathcal{E}^{(i)}\} - p(\mathcal{E})\right].$$

The last brackets, in virtue of the stationarity of the sequence $\{\eta_n\}$ with respect to the measure p, are equal to zero. Hence we get $P(\mathcal{E}^{(1)}) = P(\mathcal{E})$ for all events $\mathcal{E}$, that is, the measure P is indeed invariant.

At the end of this section we give the law for forming statistical mixtures of the measures P and p.

Theorem 2. If a measure p_i corresponds, in the preceding sense, to a measure P_i, then the measure

$$(\sum_i \alpha_i\lambda_i)^{-1} \cdot \sum_i \alpha_i\lambda_i p_i \tag{10}$$

corresponds to the measure $\Sigma_i \lambda_i P_i$, where $\lambda_i > 0$ and $\Sigma_i \lambda_i = 1$.

The proof of this theorem is not difficult. This rule has, however, an important consequence which is not quite evident.

Theorem 3. The sequence $\{\xi_i\}$ is metrically transitive with respect to the measure P if and only if the sequence $\{\eta_i\}$ has the same property with respect to p.

By mëtrically transitive we mean that each event concerning the variables ξ_i which is invariant, under the transformation $\{\xi_i\} \longrightarrow \{\xi_{i+1}\}$, has P-probability equal to zero or one.

Proof. The set of P-measures and the set of p-measures are convex. From theorem 2 it follows that the extremal points of one of the convex sets are mapped onto the extremal points of the other. On the other hand, extremality and transitivity are equivalent, which concludes the proof.

This elegant method, based on notions and theorems of the theory of convex sets, was successfully used in ergodic theory by Savage and Hewitt.

3. The recurrence transformation

We shall consider a probability space (Ω, B, μ) with a measure preserving transformation T, that is, a point transformation from the space Ω into itself, satisfying the following conditions: $T^{-1}(B) \subset B$ and $\mu(T^{-1}(\mathcal{E})) = \mu(\mathcal{E})$ for all $\mathcal{E} \in B$. We shall say that T is 1 to 1 measure preserving if it is 1 to 1 from the space Ω onto itself, if $T(B) = B$ and if T is measure preserving. Let $\mathcal{E}$ be a fixed measurable set of positive measure. By the famous Poincaré-Zermelo theorem for almost every point $\omega \in \mathcal{E}$ some of its iteration $T^k(\omega)$ for $k \geqq 1$ also belongs to the set $\mathcal{E}$. More precisely, we have the relation

$$(11) \qquad \lim_{n} \frac{1}{n} \sum_{k=1}^{n} \chi_{\mathcal{E}}(T^k\omega) > 0 \quad \text{for a.e. } \omega \in \mathcal{E},$$

where $\chi_{\mathcal{E}}$ denotes the characteristic function of the set $\mathcal{E}$. Hence we can define on the set $\mathcal{E}$ the recurrence transformation $T_{\mathcal{E}}$ by condition

$$(12) \qquad T_{\mathcal{E}}(\omega) \overset{df}{=} T^k(\omega) \in \mathcal{E} \quad \text{and} \quad T^i(\omega) \notin \mathcal{E} \quad \text{for } 1 \leqq i < k.$$

More exactly, this transformation $T_{\mathcal{E}}$ must be considered on the set

$$(13) \qquad \mathcal{E}_0 \overset{df}{=} \{\omega\colon \omega \in \mathcal{E} \text{ and } \overline{\lim_{k}}\, \chi_{\mathcal{E}}(T^k\omega) = 1\}.$$

Now it is easy to verify that $\mathcal{E}$ and $\mathcal{E}_0$ are almost equal. In our probabilistic considerations, sets of measure zero may be neglected. We emphasize that the recurrence transformation depends on the choice of the set $\mathcal{E}$ and is defined only on it.

The basic property of the recurrence transformation is

Theorem 4. $T_{\mathcal{E}}$ preserves the measure μ in the new measure space $(\mathcal{E}, \mathcal{E} \cap B, \mu)$.

In probabilistic applications we can also consider the normed conditional measure $\mu_{\mathcal{E}}(\cdot) \overset{\mathrm{df}}{=} \mu(\cdot|\mathcal{E})$. For the proof it suffices to verify that

$$(14) \qquad \mu(\mathcal{X}) = \sum_{k=1}^{\infty} \mu[\mathcal{E} \cap T^{-1}(\mathcal{E}^c) \cap \cdots \cap T^{-k+1}(\mathcal{E}) \cap T_{\mathcal{E}}^{-k}(\mathcal{X})].$$

It is easy to see that the right side of formula (14) is equal to $\mu[T_{\mathcal{E}}^{-1}(\mathcal{X})]$.

On the other hand, we obtain by a simple computation

$$(15) \qquad S_n = \sum_{k=1}^{n} \mu[\mathcal{E} \cap T^{-1}(\mathcal{E}^c) \cap \cdots \cap T^{-k+1}(\mathcal{E}^c) \cap T^{-k}(\mathcal{X})]$$

$$= \mu\,[F_n \cap T^{-n}(\mathcal{X})],$$

where $F_n = \mathcal{E} \cup \cdots \cup T^{-n}(\mathcal{E})$. We observe now that the limit set $\lim_n F_n = F = \bigcup_{n=0}^{\infty} T^{-n}(\mathcal{E})$ is almost T-invariant and contains the set $\mathcal{X}$. Hence we can write

$$(16) \qquad S_n = \mu[F \cap T^{-n}(\mathcal{X})] - \mu\,[(F-F_n) \cap T^{-n}(\mathcal{X})]$$

$$= \mu(\mathcal{X}) - \mu[(F-F_n) \cap T^{-n}(\mathcal{X})].$$

To conclude the proof it suffices to remark that the second term on the right side converges to zero.

It is easy to observe close connection of the recurrence transformations with the theory of stationary sequences of calls. For this purpose it is enough to consider the exponent k, in the definition (5), as a function of a point $\omega \in \mathcal{E}_0$. The measurability of this function $k = k(\omega)$ is clear. Now we form a sequence

(17) $$k(\omega), k(T_{\mathcal{E}}\omega), k(T_{\mathcal{E}}^2\omega), \cdots .$$

In view of theorem 4 this sequence is stationary with respect to the measure $\mu_{\mathcal{E}}$, and has exactly the same meaning as the random variables $\eta_0, \eta_1, \eta_2, \cdots$ considered above. We must put $\mathcal{E} \overset{df}{=} \{\xi_0 = 1\}$, $T \overset{df}{=}$ the shift transformation and we have $\mu \overset{df}{=} P$ and $\mu_{\mathcal{E}} = p$.

In addition we remark that for each 1 to 1 measure preserving transformation T the transformation $T_{\mathcal{E}}$ is also 1 to 1, and therefore we can also form the negative iterations

(18) $$\cdots, k(T^{-3}\omega), k(T^{-2}\omega), k(T^{-1}\omega).$$

We are not going to give a systematic study of the recurrence transformations. We shall present some formulas and properties only.

(i) $T_{\mathcal{E}_1} = (T_{\mathcal{E}_2})_{\mathcal{E}_1}$ for measurable sets $\mathcal{E}_1 \subset \mathcal{E}_2$.

(ii) If the transformation T is metrically transitive then $T_{\mathcal{E}}$ is also transitive.

(iii) If we suppose in addition that the iterations of $\mathcal{E}$ cover the whole space Ω, then the inverse implication is also true.

The proof of (i) consists of a simple calculation based on the definition (12). The proof of (ii) and (iii) is the same as that of theorem 3.

We can raise different problems about the recurrence transformations. Examples are various mixing properties hereditary from T on $T_{\mathcal{E}}$.

4. The conditional probability for arbitrary processes of calls

We start with a precise description of the measurable space for the process of calls. Let Ω denote the class of all countable subsets of the real axis R, which is the time axis, that are finite in each finite interval. The elements of Ω, which are the realizations of our process, are denoted by ω. By $N(\omega, Q)$, or simply by $N(Q)$ we denote the number of calls in the time set Q, that is,

(19) $$N(\omega, Q) \overset{df}{=} \text{card } (\omega \cap Q).$$

Now we define the class B of measurable subsets of Ω as the σ-field generated by all the functions $N(\cdot, Q)$, where Q is a Borel set, that is, B is the smallest σ-field with respect to which all functions $N(\cdot, Q)$ are measurable. It is obvious that in the preceding definition we can replace the family $\{Q\}$ of all Borel sets by the family of all intervals, or by the family of the intervals with rational endpoints. Evidently each ω can also be treated as a purely atomic measure, finite for bounded sets. We emphasize that (Ω, B) has good set theoretic structure. Namely it is not difficult to prove that Ω can be mapped by a 1 to 1 function onto the unit interval I and the class B onto the class of all I Borel subsets. Hence the typical difficulties of the theory of conditional probability do not occur in our space (Ω, B).

Let a fixed probability measure P be defined on (Ω, B). Our next aim is to give a precise meaning to the notion: the probability of an event $\mathcal{E}$ under the condition that a call occurred at moment t. For this notion, not yet defined, we shall use the symbols $P(\mathcal{E}|t)$ or $P(\mathcal{E}|t \in \omega)$. Now, we introduce a new assumption, which is quite natural and at the same time seems to be necessary for our consideration. We suppose that

$$\int N(\omega,Q)P(d\omega) = E_P[N(Q)] < \infty \tag{20}$$

for all bounded sets Q. Consequently we put $\mu(Q) \overset{df}{=} E_P[N(Q)]$. Obviously μ is a Borel measure on the time axis.

For each event $\mathcal{E} \in B$ the integral

$$\int \chi_{\mathcal{E}}(\omega)N(\omega,Q)P(d\omega), \tag{21}$$

treated as a function of the set Q, is an absolutely continuous measure with respect to μ. Hence by the Radon-Nikodym theorem we can write

$$\int \chi_{\mathcal{E}}(\omega)N(\omega,Q)P(d\omega) = \int_Q P(\mathcal{E}|t)\mu(dt). \tag{22}$$

For each fixed $\mathcal{E}$ the Radon-Nikodym derivative $P(\mathcal{E}|t)$ is unique only a.e. with respect to μ, and we can always suppose that it is a "true" measure with respect to sets $\mathcal{E} \in B$. This follows from the previously mentioned property that a measurable space (Ω, B) is a Borel space.

Formula (20) can be generalized to

$$\int f(\omega)N(\omega,Q)P(d\omega) = \int_Q \mu(dt)\int f(\omega)P(d\omega), \tag{23}$$

where f is a P-integrable function.

It seems that this way of introducing the probability $P(\mathcal{E}|t)$ is appropriate. We shall only remark that

(i) If there is exactly one call then $P(\mathcal{E}|t)$ is identical with the ordinary conditional probability.

(ii) If some process of calls is realized by the sequence of random variables $x_1, x_2, \cdots,$ for which $\Pr\{x_i \neq x_j\} = 1$ for $i \neq j$, in the following sense $N(Q) \overset{df}{=} \Sigma_i \chi_Q(x_i)$ for all Borel sets Q, then our assumption (20) takes on the form $\Sigma_i \mu_i(Q) = \mu(Q) < \infty$ for all bounded Q, where $\mu_i(Q) \overset{df}{=} \Pr\{x_i \in Q\}$. Moreover, the probability $P(\mathcal{E}|t)$ can be written

$$P(\mathcal{E}|t) = \sum_i P(\mathcal{E}|x_i = t) P(x_i = t|t), \tag{24}$$

where $P(x_i = t|t)$ is equal to the Radon-Nikodym derivative $d\mu_i/d\mu$ and can be interpreted as the conditional probability that the ith call occurs at the moment t, given that a call appears at this moment.

5. The stationary process

Now we shall consider stationary processes of calls. We shall use the following notations for shifts

$$\begin{aligned} &\omega^t \overset{df}{=} \omega + t \text{ for } \omega \in \Omega \text{ and } -\infty < t < \infty, \\ &\omega \in \mathcal{E}^t \overset{df}{=} \omega^{-t} \in \mathcal{E}. \end{aligned} \tag{25}$$

We add the new assumption $P(\mathcal{E}) = P(\mathcal{E}^t)$ for all $\mathcal{E} \in B$ and $-\infty < t < \infty$. As in the case of the discrete time, we have

$$P\{N(-\infty, +\infty) = 0 \text{ or } N(-\infty, 0) = N(0, \infty) = \infty\} = 1, \tag{26}$$

and in what follows we always assume that

(27) $$P\{N(-\infty, +\infty) = 0\} = 0.$$

Hence we can restrict our attention to the realizations ω with infinitely many calls in both time directions. First we shall establish the form of the conditional probability $P(\mathcal{E}|t)$, from the preceding section.

Theorem 5. There exists one and only one probability measure P_0 defined on the space (Ω, B) for which the measure function

(28) $$\Pi(\mathcal{E}|t) \overset{df}{=} P_0(\mathcal{E}^{-t}),$$

depending on the parameter t, satisfies the condition (22) for all $\mathcal{E} \in B$ and Q.

Proof. For the stationary process the measure μ is proportional to the Lebesgue measure $\mu(dt) = \alpha dt$, where α is the intensity of calls. Roughly speaking, the matter is quite simple: the new measure $P_0(\mathcal{E})$ is equal to $P(\mathcal{E}|0)$ and formula (28) is a special case of

(29) $$P(\mathcal{E}|t) = P(\mathcal{E}^s|t+s),$$

which seems to be obvious in view of the assumed stationarity. For a precise proof, let $P(\mathcal{E}|t)$ be any conditional probability measure satisfying (22). It follows from the stationarity that for each $\mathcal{E} \in B$, $s \in R$ and for almost every t, in the sense of the Lebesgue measure, the relation (29) holds. In view of the Fubini theorem we know that there is a number t_0 such that, for almost every s and for all $\mathcal{E} \in B$ we have $P(\mathcal{E}|t_0) = P(\mathcal{E}^s|t_0 + s)$.

The quantifier "for all $\mathcal{E}$" can be put at the end because $P(\cdot|t)$ is a measure and it suffices to consider only some countable class of sets $\mathcal{E}$ generating the whole field B. It is easy to see that the measure $P_0(\mathcal{E}) \overset{df}{=} P(\mathcal{E}^{t_0}|t_0)$ satisfies the assertion of theorem 5. The proof of the uniqueness of the measure P_0 is omitted because it is quite simple.

Remark 1. Now our conditional probability measure $P(\cdot|t)$ can be determined in a unique manner by equation (29). This "regular" $P(\cdot|t)$ will be used later.

Remark 2. Now the equation (23) takes the form

$$(30) \qquad \int f(\omega)N(\omega,Q)P(d\omega) = \alpha\int P_0(d\omega)\int_Q f(\omega^t)dt.$$

Next we give another description of P_0. We say that a B-measurable function $f(\omega)$ defined on Ω is continuous if and only if it is bounded and if for each fixed $\omega \in \Omega$ the function $f(\omega^t)$ of the real variable t is continuous.

Theorem 6. *If f is continuous, then*

$$(31) \qquad \lim_{|I|\to 0} \frac{1}{\alpha|I|}\int_{\{N(I)\geq 1\}} f(\omega)P(d\omega) = \int f(\omega)P_0(d\omega),$$

where $0 \in I$, an interval. Or, in another form,

$$(32) \qquad \lim_{|I|\to 0} E_P(f|N(I) \geq 1) = EP_0(f).$$

Proof. From formula (30) we have

$$(33) \qquad \lim_{|I|\to 0} \frac{1}{\alpha|I|}\int f(\omega)N(\omega,I)P(d\omega)$$

$$= \int P_0(d\omega) \lim_{|I|\to 0}\frac{1}{|I|}\int f(\omega^t)dt = \int P_0(d\omega)f(\omega),$$

and on the other hand

$$(34)\qquad \left|\frac{1}{\alpha|I|}\int f(\omega)N(\omega,I)P(d\omega) - \frac{1}{\alpha|I|}\int_{\{N(I)\geq 1\}} f(\omega)P(d\omega)\right|$$

$$\leq \sup_{\omega\in\Omega} |f(\omega)| \frac{1}{\alpha|I|}\int_{\{N(I)\geq 1\}} [N(\omega,I)-1]P(d\omega).$$

The right side tends to zero together with the length of I (compare the theorem of Koroliuk, p. 39 of [2]).

Corollary. $P_0(\Omega_0) = 1$, where $\Omega_0 \overset{df}{=} \{\omega\colon 0 \in \omega\}$.

For the proof of the preceding very intuitive equality, we consider a sequence $\{f_n(\omega)\}$ of continuous functions defined as

$$(35)\qquad f_n(\omega) = \begin{cases} 0 & \text{if } d(\omega) \geq \frac{1}{n} \\ 1-nd(\omega) & \text{if } d(\omega) < \frac{1}{n}, \end{cases}$$

where $d(\omega)$ is the distance of the set ω from the point $t=0$. By theorem 6 we have $\int f_n(\omega)P_0(d\omega) = 1$ for $n=1,2,\cdots,$ and hence in the limit we obtain $P_0(\Omega_0) = 1$.

We introduce now a sequence $\{\eta_n(\omega)\}$ of functions defined on the subspace Ω_0 sketched in Figure 2.

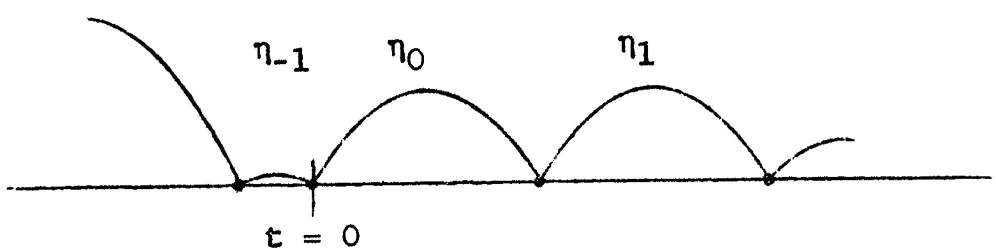

Figure 2

Theorem 7. The random variables $\{\eta_n\}$ $n=0,\pm1,\pm2,\cdots$ in the measure space $(\Omega_0, \Omega_0 \cap B, P_0)$ are

(i) nonnegative $\eta_n > 0$,

(ii) they have finite first moment $E_{P_0}(\eta_n) = \alpha^{-1} < \infty$,

(iii) $\eta_1, \eta_2, \cdots$ form a stationary sequence with respect to the measure P_0.

Conversely each sequence of random variables $\{\eta_n\}$ satisfying (i) to (iii) can be obtained in this way. Moreover, the correspondence between P and P_0 is 1 to 1 and it is given by

$$\int f(\omega)P(d\omega) = \alpha \int_{\Omega_0} P_0(d\omega) \int_0^{\eta_0} f(\omega^{-t})dt, \tag{36}$$

which is valid for all measurable and P-integrable function f.

Proof. Let a be a positive number. We have, in virtue of the stationarity,

$$\int_{\{N(-a,0)>0\}} f(\omega)P(d\omega) = \sum_{k=1}^{n} \int_{\{N(I_{-k})>0, N(I_{-k+1})=0, \cdots, N(I_{-1})=0\}} f(\omega)P(d\omega) \tag{37}$$

$$= \frac{1}{\delta} \sum_{k=1}^{n} \int_{\{N(I_{-1})>0, N(I_0)=0, \cdots, N(I_{k-2})=0\}} f(\omega^{-(k-1)\delta})P(d\omega),$$

where $\delta \overset{df}{=} a/n$ and $I_k \overset{df}{=} [k\delta, (k+1)\delta)]$. Consequently, we obtain

$$\int_{\{N(-a,0)\ 0\}} f(\omega)P(d\omega) = \frac{\alpha}{\alpha\delta} \int P(d\omega)\delta \sum_{k=1}^{m_n(\omega)} f(\omega^{-(k-1)\delta}), \tag{38}$$

where $m_n \overset{df}{=} n[\varphi/n] - 1$ and $\varphi \overset{df}{=} \min(\eta_0, a)$. When $n \longrightarrow \infty$, the Riemann sums

$$\delta \sum_{k=1}^{m_n} f[\omega^{-(k-1)\delta}] \to \int_0^{\varphi} f(\omega^{-t})dt \tag{39}$$

in a bounded manner with respect to the variable ω. In virtue of theorem 6 we obtain

(40) $$\int_{\{N(-a,0)>0\}} f(\omega)P(d\omega) = \alpha\int P_0(d\omega)\int_0^{\min(\eta_0,a)} f(\omega^{-t})dt.$$

Finally, if $a \longrightarrow +\infty$ in the last formula, we obtain the equality (36) for each continuous f and, consequently, for each P-integrable function.

Now from formula (36) we obtain, by putting $f \equiv 1$,

(41) $$1 = \alpha\int_{\Omega_0}\eta_0 P_0(d\omega) \quad \text{and} \quad E_{P_0}(\eta_0) = \alpha^{-1}.$$

Next we must prove property (iii). Let $\chi_\delta(\omega)$ be the characteristic function of the event $\{\omega: N(\omega,I) \geq 1\}$, where $I = \langle 0, \delta \rangle 0$.

From (31) and (36) we have for each continuous f

(42) $$\int f(\omega)P_0(d\omega) = \lim_{\delta\to 0} \frac{1}{\alpha\delta}\int f(\omega)\chi_\delta(\omega)P(d\omega)$$

$$= \lim_{\delta\to 0}\int_{\Omega_0} P_0(d\omega)\frac{1}{\delta}\int_0^{\eta_0} f(\omega^{-t})\chi_\delta(\omega^{-t})dt$$

$$= \lim_{\delta\to 0}\int_{\Omega_0} P_0(d\omega)\frac{1}{\delta}\int_{\max(\eta_0-\delta,0)}^{\eta_0} f(\omega^{-t})dt$$

$$= \int_{\Omega_0} f(\omega^{-\eta_0})P_0(d\omega).$$

Hence we have obtained the important equality

(43) $$\int_{\Omega_0} f(\omega)P_0(d\omega) = \int_{\Omega_0} f(\omega^{-\eta_0})P_0(d\omega),$$

valid for all continuous f, and consequently for all bounded measurable functions since, by iterated passage to the limit, we obtain all measurable functions from continuous functions. The last formula expresses the stationarity of $\{\eta_n\}$. Hence statements (i) to (iii) are proved.

We shall now give the proof of the inverse implication. We suppose that (i) to (iii) hold and a probability measure P_0 satisfies (43). We define the measure P by formula (36).

$$\int f(\omega)P(d\omega) \overset{df}{=} \beta \int_{\Omega_0} P_0(d\omega) \int_0^{\eta_0} f(\omega^{-t})dt, \tag{44}$$

where

$$\beta \overset{df}{=} \left[\int_{\Omega_0} \eta_0(\omega)P_0(d\omega)\right]^{-1}. \tag{45}$$

From (43) we obtain

$$\int f(\omega)P(d\omega) = \beta \int_{\Omega_0} P_0(d\omega) \int_{\eta_0+\cdots\rho_{k-1}}^{\eta_0+\cdots+\eta_k} f(\omega^{-t})dt \tag{46}$$

for $k=1,2,\cdots$ and consequently,

$$\int f(\omega)P(d\omega) = \beta \lim_{k\to\infty} \int_{\Omega_0} P_0(d\omega)\frac{1}{k+1} \int_0^{\eta_0+\cdots+\eta_k} f(\omega^{-t})dt, \tag{47}$$

and by analogy for arbitrary real s

$$\int f(\omega^{-s})P(d\omega) = \beta \lim_{k\to\infty} \int_{\Omega_0} P_0(d\omega)\frac{1}{k+1} \int_s^{\eta_0+\cdots+\eta_k+s} f(\omega^{-t})dt. \tag{48}$$

Hence

$$\int f(\omega^{-s})P(d\omega) = \int f(\omega)P(d\omega), \tag{49}$$

that is, the measure P is invariant.

In addition we must prove that the conditional probability induced by P and denoted for the moment by P^* is identical with P_0. We have by (36)

$$\int f(\omega)P(d\omega) = \alpha\int P^*(d\omega) \int_0^{\eta_0} f(\omega^{-t})dt. \tag{50}$$

Let χ_δ have the same meaning as before. We obtain

$$\int \chi_\delta(\omega)P(d\omega) = \beta \int_{\Omega_0} P_0(d\omega)\min(\delta, \eta_0), \tag{51}$$

and

$$\lim_{\delta\to 0} \frac{1}{\delta}\int \chi_\delta(\omega)P(d\omega) = \alpha, \quad \lim_{\delta\to 0} \frac{1}{\delta} \beta \int_{\Omega_0} P_0(d\omega)\min(\delta,\eta_0) = \beta. \tag{52}$$

Therefore, $\alpha = \beta$.

From (44) we have, by the previous reasoning, the equality

$$\int P^*(d\omega)f(\omega) = \int P_0(d\omega)f(\omega^{-\eta_0}) \tag{53}$$

which is valid for all continuous f. Then $P^* = P_0$.

Finally we give an analogy of theorem 3 which can be proved in a similar way.

Theorem 8. The measure P is metrically transitive if and only if the measure P_0 is metrically transitive.

REFERENCES

[1] C. PALM, "Intensitätsschwankungen im Fernsprechverkehr, Ericsson Technics, Vol. 44 (1943), pp. 1-89.

[2] A. J. KHINCHINE, Mathematical Methods of Servicing Problems, Moscow, 1955.

Statistical estimation of semantic provability.

Antonín Špaček.

Let us point out that there is nothing unexpected in this paper. The sole element of novelty is the formal description of a simple relation between a chapter of mathematical logic and mathematical statistics. The word semantic occurring in the title indicates that, roughly speaking, provability or non-provability is to be estimated on the basis of truth and falsehood in interpretations in models. The logical formalism used in this paper is monadic logic introduced by P.R. Halmos in [2]. In principle it is possible to replace the monadic logic by a more developed formalism, for instance, by polyadic logic [3]. The elements, the provability or non-provability of which is to be estimated, as well as the interpretations are chosen at random by appropriate chance mechanisms, hence, the whole problem is probabilistic in nature. The estimation procedures established in this paper possess a natural optimum property. The study of behaveor of these procedures at infinity shows that the statistical decision functions of finite size, which estimate provability are, in fact, asymptotically good approximate proofs. One may hope that the questions treated in this paper reflect at least the most elementary features of heuristic reasoning which is so perfectly realized by the human brain.

All that is necessary for an easy understanding is developed in the paper in full details and with intuitive justification. The main reason is that one cannot expect that, in general, specialists in mathematical logic are familiar with concepts, methods and results of statistical decision theory and that statisticians are familiar with formalisms of mathematical logic.

The basic concepts and results of statistical decision theory on an appropriate level of generality are summarized in §1. These results are then applied in §2 to the problem of statistical estimation of belonging relations. The passage from the considerations of §2 to the solution of our main problem of statistical estimation of provability is completely transparent and forms the contents of §3.

The present paper, which is closely connected with [7], does not furnish more than may be intuitively expected and, therefore, its practical value is very limited. Further developments in this direction, however, will probably throw some light into the mechanism of human behaveor in problem solving.

§1. The Neyman - Pearson theorem.

A wide variety of problems of mathematical statistics can be reduced to a simple application of a classical theorem due to J.Neyman and E.S.Pearson [5]. It is not surprising that this famous theorem plays a decisive role in our considerations.

Its original version, however, does not fulfil our requirements. The main reason is that it does not allow the discussion of cases in which more general sample spaces occur. We shall see later that an adequate generalization of the Neyman-Pearson theorem can be easily obtained.

Our basic probability space will be denoted, as usual, by $(\Omega, \mathfrak{S}, \mu)$, where Ω is the set of elementary events, $\mathfrak{S}$ the sigma-algebra of random events and μ the probability measure on $\mathfrak{S}$. The symbol ω will always mean an element of Ω. Throughout this paper the notation just introduced will be preserved.

A statistical decision problem is defined to be a pair (φ, ξ) of random variables, where φ takes its values in the parameter space and ξ ranges over the sample space. The parameter space is assumed to consist of exactly two elements, namely 0 and 1, hence, the measurability of φ is assured by the requirement that

$$\{\omega : \varphi(\omega) = 1\} \in \mathfrak{S}$$

On the other hand, no restriction will be imposed on the range X of ξ except that it is supplied with a fixed sigma-algebra $\mathfrak{X}$ of subsets of X. The measurable space $(X, \mathfrak{X})$ is said to be the sample space. The transformation ξ of Ω into X will be called random sample if

$$\{\omega : \xi(\omega) \in E\} \in \mathfrak{S}$$

for every set E from $\mathfrak{X}$.

Roughly speaking, a statistical decision is an action determined by the value of the random sample. This action can be formally described using the concept of decision function. The domain of a decision function is the sample space and its range is usually called the space of decisions. In our case, however, the space of decisions is assumed to coincide with the parameter space, hence, a decision function δ is a function defined on X and taking the values 0 or 1. But this is not enough. In order to ensure that the compound transformation $\delta(\xi(.))$ becomes a random variable, it is reasonable to impose on δ an additional condition of measurability, namely,

$$\{x : \delta(x) = 1\} \in \mathfrak{X} .$$

A natural manner of how to evaluate statistical decisions with respect to the random occurence of parameters is the convention that

$$(\{\omega : \varphi(\omega) = 1\} \cap \{\omega : \delta(\xi(\omega)) = 0\}) \cup (\{\omega : \varphi(\omega) = 0\} \cap \{\omega : \delta(\xi(\omega)) = 1\})$$

means the random event of incorrect decisions.

Our main question is how to choose the decision function δ in order to make the probability of the random event of incorrect decisions as small as possible. The answer is quite satisfactory.

THEOREM 1. There always exists a statistical decision function which minimizes the probability of the random event of incorrect decisions.

The proof is a simple application of the Hahn decomposition theorem 4 . Let us write

$$\nu(E) = \mu(\{\omega : \varphi(\omega) = 1\} \cap \xi^{-1}(E)) - \mu(\{\omega : \varphi(\omega) = 0\} \cap \xi^{-1}(E))$$

for every E from $\mathcal{H}$. Clearly, ν is a signed measure on $\mathcal{H}$. It is well known that there exists a set H from $\mathcal{H}$ such that $\nu(H\cap E)\geqslant 0$ and $\nu(H'\cap E)\leqslant 0$ for every E from $\mathcal{H}$, where $H'=X-H$. Since

$$\nu(E)=\nu(H\cap E)+\nu(H'\cap E)\leqslant \nu(H)+\nu(H'\cap E)\leqslant \nu(H)$$

for every E from $\mathcal{H}$, hence, the number $\nu(H)$ is the maximum of ν on $\mathcal{H}$. Now let us define the decision function β by the requirement that

$$\{x:\beta(x)=1\}=H.$$

Since for every decision function δ the probability of the random event of incorrect decisions is equal to

$$\mu\{\omega:\varphi(\omega)=1\}-\nu\{x:\delta(x)=1\}$$

hence, using the fact that β is determined by the Hahn decomposition (H, H') of ν, we can write

$$\mu\{\omega:\varphi(\omega)=1\}-\nu\{x:\beta(x)=1\}\leqslant \mu\{\omega:\varphi(\omega)=1\}-\nu\{x:\delta(x)=1\}$$

for every decision function δ, q.e.d.

The decision function β, the existence of which is assured by Theorem 1, is said to be the Bayes solution of the statistical decision problem (φ, ξ).

It is easy to verify that the signed measure ν is absolutely continuous with respect to the probability measure $\mu\xi^{-1}$ in $\mathcal{H}$, hence, using the Radon-Nikodym theorem [4], we can state that there exists a real valued measurable function h on X such that

$$\nu(E)=\int_E h(x)\,d\mu\xi^{-1}$$

for every set E from $\mathcal{X}$. We see at once that the set

$$\{x : h(x) > 0\}$$

and its complement determine a Hahn decomposition of ν and this is in fact the content of the Neyman-Pearson theorem. It is, however, more appropriate to formulate this theorem in terms of the measurable functions h^+ and h^- defined for every element x of X and every set E from $\mathcal{X}$ by the equations

$$\mu(\{\omega : \varphi(\omega) = 1\} \cap \xi^{-1}(E)) = \alpha \int_E h^+(x)\, d\mu\xi^{-1},$$

$$\mu(\{\omega : \varphi(\omega) = 0\} \cap \xi^{-1}(E) = (1-\alpha) \int_E h^-(x)\, d\mu\xi^{-1},$$

where

$$\alpha = \mu\{\omega : \varphi(\omega) = 1\}$$

The number α is said to be the a priori probability in the parameter space. Clearly, if $\alpha > 0$ then h^+ is a conditional probability density and if $\alpha < 1$ then h^- is a conditional probability density. Since

$$\mu\xi^{-1}\{x : h(x) = \alpha h^+(x) - (1-\alpha) h^-(x)\} = 1,$$

the Neyman-Pearson theorem can be formulated as follows:

THEOREM 2. The statistical decision function β determined by the relation

$$\{x : \beta(x) = 1\} = \{x : \alpha h^+(x) > (1-\alpha) h^-(x)\}$$

minimizes the probability of the random event of incorrect decisions.

In applications of this theorem the densities h^+ and h^- are always assumed to be known, hence, the Bayes solution β of the statistical decision problem (φ, ξ) depends only on the a priori probability α in the parameter space.

Now we shall introduce the abstract substitute of the concept of sample size. The classical model shows that one of the most important consequences of the reduction of sample size is a restriction imposed on the measurability of the decision functions. This fact motivates our definition of the size of a decision function.

Let $\mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_3, \ldots$ be a non-decreasing sequence of sigma-algebras of subsets of X and suppose that the union

$$\bigcup_{n=1}^{\infty} \mathcal{X}_n$$

is a base of the sigma-algebra $\mathcal{X}$. This sequence will serve as a scale of the sizes of decision functions.

The decision function δ is said to be of size n if it is measurable with respect to the sigma-algebra $\mathcal{X}_n$, i.e.

$$\{x : \delta(x) = 1\} \in \mathcal{X}_n ,$$

but it is not measurable with respect to the sigma-algebra $\mathcal{X}_m$ for $m = 1, 2, \ldots, n-1$. We shall say that δ is of finite size if there exists a positive integer n such that δ is of size n. The decision function δ, which is by definition measurable with respect to the whole sigma-algebra $\mathcal{X}$, is said to be of infinite size if it is not of finite size. Clearly, if there exists a decision function of finite size then, roughly speaking, the scale $\mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_3, \ldots$ must have effectively an infinite number of divisions.

Denoting by Δ the set of all decision functions in X and by Δ_n that of all decision functions in X at most of size n for $n = 1, 2, 3, \ldots$, we see at once that

$$\Delta_1 \subset \Delta_2 \subset \Delta_3 \subset \cdots \subset \Delta \ ,$$

hence, if ε is the probability of the random event of incorrect decisions associated with the Bayes decision function β from Δ and ε_n that associated with the Bayes decision function β_n from Δ_n for $n = 1, 2, 3, \ldots$ then

$$\varepsilon_1 \geqslant \varepsilon_2 \geqslant \varepsilon_3 \geqslant \cdots \geqslant \varepsilon$$

i.e., as may be intuitively expected, the least probabilities of making incorrect decisions do not increase whenever the sizes of the decision functions admitted to the concurrence increase to infinity.

By Theorem 2 a Bayes decision function β_n of size n is determined by the relation

$$\{x : \beta_n(x) = 1\} = \{x : \alpha h_n^+(x) > (1-\alpha) h_n^-(x)\}$$

for $n = 1, 2, 3, \ldots$, where h_n^+ and h_n^- are defined using the sigma-algebra $\mathcal{H}_n$ in the same way as h^+ and h^- were defined using the whole sigma-algebra $\mathcal{H}$.

The main effect of increasing the sample size can be expressed as follows:

<u>THEOREM 3.</u> <u>The sequence of random variables</u> $h_1^+(\xi(.))$, $h_2^+(\xi(.))$, $h_3^+(\xi(.))$, ... <u>converges to the random variable</u> $h^+(\xi(.))$ <u>with probability one</u>, <u>the sequence of random variables</u> $h_1^-(\xi(.))$, $h_2^-(\xi(.))$, $h_3^-(\xi(.))$, ... <u>converges to the random variable</u> $h^-(\xi(.))$ <u>with probability one</u>, <u>and the sequence</u> $\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots$ <u>of probabilities of the random event of incorrect decisions</u>, <u>associated successively with the Bayes decision functions</u> $\beta_1, \beta_2, \beta_3, \ldots$.

converges to the probability ε of the random event of incorrect decisions associated with the Bayes decision function β.

The first two assertions of Theorem 3 are immediate consequences of a well known martingale theorem [1] and the last assertion is contained in [6] as a particular case.

Let us note that if $\varepsilon = 0$ then the last assertion of Theorem 3 expresses the well known property of consistency of the Bayes decision functions $\beta_1, \beta_2, \beta_3, \ldots$

§2. Statistical estimation of belonging relations.

A wide variety of questions concerning statistical estimation of provability possesses a common statistical structure of very elementary nature and this fact enables us to treat the basic statistical problem separately and independently of any consideration belonging purely to the domain of mathematical logic. After establishing the general results it remains only to interpret them appropriately in order to obtain the desired final answer to various questions of statistical estimation of provability. The realization of this last step is, however, rather only a routine matter.

Suppose that one wants to decide whether an element chosen at random by an appropriate chance mechanism from a fixed set A belongs or does not belong to a fixed non-empty proper subset M of A.

The random variable η taking values in A is assumed to be a formal substitute of our basic chance mechanism. One of the most natural requirements concerning measurability is

$$\{\omega : \eta(\omega) \in M\} \in \mathfrak{S}.$$

The direct observation on M is replaced by observations on the subsets $Q(m)$ of A for $m = 1, 2, 3, \ldots$, hence, it is also natural to impose on η an additional condition, namely,

$$\{\omega : \eta(\omega) \in Q(m)\} \in \mathfrak{S}$$

for $m = 1, 2, 3, \ldots$ and this completes the definition of the random variable η.

Now let τ be an ordinary random variable taking on values of positive integers. The compound transformation $Q(\tau(.))$ is a random variable in the sense that

$$(1) \qquad \{\omega : p \in Q(\tau(\omega))\} \in \mathfrak{S}$$

for every element p of A. This follows from the obvious identity

$$\{\omega : p \in Q(\tau(\omega))\} = \bigcup_{j=1}^{\infty} \{\omega : \tau(\omega) = m_j\},$$

where m_j is the j-th subscript for which $p \in Q(m_j)$.

Clearly,

$$\{\omega : \eta(\omega) \in Q(\tau(\omega))\} = \bigcup_{m=1}^{\infty} (\{\omega : \eta(\omega) \in Q(m)\} \cap \{\omega : \tau(\omega) = m\}),$$

hence, we can state that

$$(2) \qquad \{\omega : \eta(\omega) \in Q(\tau(\omega))\} \in \mathfrak{S}$$

and this is the most important fact concerning the relation between the two kinds of random variables.

In elementary set theory the relation $p \in M$ is often expressed in terms of the characteristic function c of M by the equivalent statement that $c(p) = 1$. A little more complicated concept is that of the characteristic function of a random set. If for each $m = 1, 2, 3, \ldots$ $c(m)$ denotes the characteristic function of the set $Q(m)$ then by (1) the compound function $c(\tau(.))$ is an ordinary random variable taking the values 1 or 0. The random variable $c(\tau(.))$ is said to be the characteristic function of the random set $Q(\tau(.))$. The element p of A belongs to $Q(\tau(\omega))$ or to its complement $A - Q(\tau(\omega))$ according as $c(\tau(\omega))(p) = 1$ or $c(\tau(\omega))(p) = 0$.

Clearly, the compound transformation $c(\eta(.))$ is an ordinary random variable taking the values 1 or 0. The value of η at ω belongs to M or to its complement $A-M$ according as $c(\eta(\omega)) = 1$ or $c(\eta(\omega)) = 0$. By (2) the compound transformation $c(\tau(.))(\eta(.))$ is an ordinary random variable taking the values 1 or 0. The value of η at ω belongs to $Q(\tau(\omega))$ or to its complement $A - Q(\tau(\omega))$ according as $c(\tau(\omega))(\eta(\omega)) = 1$ or $c(\tau(\omega))(\eta(\omega)) = 0$. We have thus defined a probabilistic extension of belonging relations.

In order to simplify the notation we shall write $\varphi(.)$ instead of $c(\eta(.))$ and $\chi(.)$ instead of $c(\tau(.))(\eta(.))$.

Let X be the set of all sequences $x = (x_1, x_2, x_3, \ldots)$ every term of which is either equal to 1 or to 0. The coincidence in the first n terms of sequences from X is an equivalence relation in X. The class $\mathcal{X}_n$ of all unions of equivalence sets induced by this equivalence relation is a complete algebra of subsets of X for every $n = 1, 2, 3, \ldots$ The sets from $\mathcal{X}_n$ are called n-dimensional cylinders. Our basic sigma-algebra $\mathcal{X}$ of subsets of X is that induced by the union

$$\bigcup_{n=1}^{\infty} \mathcal{H}_n .$$

Let $\tau_1, \tau_2, \tau_3, \ldots$ be a sequence of integral-valued random variables. Then the sequence

$$\xi = (\chi_1, \chi_2, \chi_3, \ldots) ,$$

where $\chi_n(.) = c(\tau_n(.))(\eta(.))$ for $n = 1, 2, 3, \ldots$, is a random vector with values in X. Clearly, $\mathcal{H}$ is the smallest sigma-algebra of subsets of X for which the vector ξ is measurable.

Now the ground is prepared to put the traditional machinery of statistical decision functions into action. The passage from the general scheme of statistical decision to our particular case is very simple because the notation of §1 is preserved. As has already been pointed out in §1, the Bayes solution of a statistical decision problem depends on the a priori probability in the parameter space. We shall see, however, that, as compared with the general case, our particular version of the statistical decision problem is, roughly speaking, less sensitive to the exact knowledge of the a priori probability, provided that a very simple and natural condition, namely,

$$M \subset Q(m) \tag{3}$$

for $m = 1, 2, 3, \ldots$ is satisfied. We shall see that under this condition either the decision function β_n^* which associates with every sample point x of X the decision 0, or the decision function β_n which associates with the sample point x of X the decision 1 or 0 according as the first n coordinates of x are equal to 1 or at least one of these coordinates is equal to 0, can occur. More precisely:

THEOREM 4. Under (3) the Bayes solution of size n of the statistical decision problem (φ, ξ) is determined by the decision function β_n or β_n^* and the probability of the random event of incorrect decisions is equal to

$$(1-\alpha)\, h_n^-(1,1,\ldots,1,x_{n+1},x_{n+2},\ldots) \tag{4}$$

or to α according as

$$h_n^-(1,1,\ldots,1,x_{n+1},x_{n+2},\ldots) < \alpha/(1-\alpha) \tag{5}$$

or the opposite inequality holds.

The details of the proof can be omitted because Theorem 4 is nothing else but a particular version of Theorem 2. It suffices to note that, as compared with Theorem 2, the main simplification arises from (3) and from the definition of X, $\mathcal{X}_n$ and ξ. Under these conditions $h_n^+(x)=1$ or 0, according as the first n coordinates of x are equal to 1 or one at least of these coordinates is equal to 0, and $0 \leqslant h_n^-(x) \leqslant 1$ for every x from X, hence, Theorem 2 is immediately applicable.

In order to make the intuitive content of the theorem just established more transparent we shall give the informal description of an experimental procedure of how to estimate that an element of A chosen by η belongs to M or to its complement $A-M$ using the Bayes decision procedure of size n i.e. that determined by the random variables $\tau_1, \tau_2, \ldots, \tau_n$. Whenever the inequality (5) does not hold then the value of η is always estimated to belong to $A-M$. If (5) holds then the decision procedure runs as follows: At the first step we choose the set $Q(m)$ determined by the value of τ_1.

If the value of η does not belong to this set, the procedure is stopped and the value of η is estimated to belong to $A-M$. In the opposite case we continue the inspection choosing the set $Q(m)$ determined by the value of τ_2. If the value of η does not belong to this set, the procedure is stopped and the value of η is estimated to belong to $A-M$. In the opposite case we continue the inspection choosing the set $Q(m)$ determined by the value of τ_3 and so on. Exhausting all the sets $Q(m)$ determined successively by the values of $\tau_1, \tau_2, \ldots, \tau_n$ without reaching the decision that the value of η belongs to $A-M$ we accept the decision that the value of η belongs to M. We see that the final decision that the value of η belongs to $A-M$ can be reached at every step of the decision procedure. On the other hand, the opposite decision that the value of η belongs to M can be reached only at the last step.

Now we shall show that under the two additional conditions

(6) $$\bigcap_{m=1}^{\infty} Q(m) = M,$$

(7) $$\mu\Big(\bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \{\omega : \tau_n(\omega) = m\}\Big) = 1$$

the Bayes solution of the statistical decision problem (φ, ξ) becomes asymptotically independent of the a priori probability α.

Roughly speaking, the condition (6) together with (3) express the natural requirement that the approximation of M by the successive intersections of the sets $Q(m)$ can be arbitrarily close and the condition (7) means that the sequence $\tau_1, \tau_2, \tau_3, \ldots$ exhausts with probability one whole set of positive integers.

For instance, the condition (7) is satisfied whenever the integral-valued random variables $\tau_1, \tau_2, \tau_3, \ldots$ are mutually independent, identically distributed, and such that

$$\mu\{\omega : \tau_1(\omega) = m\} > 0$$

for $m = 1, 2, 3, \ldots$

Clearly, under the last condition,

$$\mu\left(\bigcap_{n=1}^{k} \{\omega : \tau_n(\omega) \neq m\}\right) = \left(\mu\{\omega : \tau_1(\omega) \neq m\}\right)^k,$$

$$\mu\{\omega : \tau_1(\omega) \neq m\} < 1,$$

for $k = 1, 2, 3, \ldots, m = 1, 2, 3, \ldots$, hence,

$$\mu\left(\bigcap_{n=1}^{\infty} \{\omega : \tau_n(\omega) \neq m\}\right) = 0$$

for $m = 1, 2, 3, \ldots$ i.e.

$$\mu\left(\bigcup_{m=1}^{\infty} \bigcap_{n=1}^{\infty} \{\omega : \tau_n(\omega) \neq m\}\right) = 0$$

or, equivalently, (7) holds.

Our Theorem 4 can be completed as follows:

THEOREM 5. If $\alpha > 0$ and the conditions (3), (6), (7) are satisfied then there exists a positive integer k such that β_n is the Bayes solution of size n of the statistical decision problem (φ, ξ) whenever $n > k$.

Since by Theorem 3 $h_n^-(\xi(\cdot)) \to h^-(\xi(\cdot))$ with probability one as $n \to \infty$, hence, by Theorem 4 and by the assumption $\alpha > 0$ of Theorem 5 it suffices to show that $h^-(1, 1, 1, \ldots) = 0$ i.e. that

(8) $$\mu\left(\{\omega : \eta(\omega) \in A - M\} \cap \{\omega : \eta(\omega) \in \bigcap_{n=1}^{\infty} Q(\tau_n(\omega))\}\right) = 0.$$

In order to simplify the notation we shall write

$$\bigcap_{m=1}^{\infty} \bigcup_{n=1} \{\omega : \tau_n(\omega) = m\} = G.$$

It follows from (6) that

$$G \cap \{\omega : \eta(\omega) \in A - M\} \cap \{\omega : \eta(\omega) \in \bigcap_{n=1}^{\infty} Q(\tau_n(\omega))\} = 0$$

hence,

$$\mu(G \cap \{\omega : \eta(\omega) \in A - M\} \cap \{\omega : \eta(\omega) \in \bigcap_{n=1}^{\infty} Q(\tau_n(\omega))\}) = 0$$

and since by (7) $\mu(G) = 1$, we obtain (8), q.e.d.

Let us denote by β the decision function which associates with every sample point x from X the decision 1 or 0 , according as $x = (1,1,1,\ldots)$ or $x \neq (1,1,1,\ldots)$. By Theorem 2, β is the Bayes solution of the statistical decision problem (φ, ξ) with respect to the whole sigma-algebra $\mathcal{X}$, hence, it is of infinite size. Since the probability of the random event of incorrect decisions is equal to zero, the decision function β , in fact, becomes a proof that the value of η belongs to M or to its complement $A - M$.

Now we shall introduce a function ℓ on Ω the values of which are positive integers or ∞ as follows:

$$\{\omega : \ell(\omega) = 1\} = \{\omega : \chi_1(\omega) = 0\},$$

$$\{\omega : \ell(\omega) = n\} = \{\omega : \chi_n(\omega) = 0\} \cap \bigcap_{j=1}^{n-1} \{\omega : \chi_j(\omega) = 1\}$$

for $n = 2, 3, 4, \ldots$ and

$$\{\omega : \ell(\omega) = \infty\} = \bigcap_{n=1}^{\infty} \{\omega : \chi_n(\omega) = 1\}.$$

We see at once that ℓ is an ordinary random variable, provided that the definition is modified in such a way that the possibility

$$\mu\{\omega : \ell(\omega) = \infty\} > 0$$

is not excluded.

The random variable ℓ is said to be the length of the decision function β.

<u>THEOREM 6. If the conditions (3), (6) and (7) are satisfied then the length of the decision function β is infinite with conditional probability one under the condition that the value of η belongs to M and it is finite with conditional probability one under the condition that the value of η belongs to $A-M$.</u>

By the definition of ℓ,

$$\mu(\{\omega: \ell(\omega)=\infty\} | \{\omega: \eta(\omega) \in M\}) =$$
$$= \frac{1}{\alpha} \mu(\{\omega: \eta(\omega) \in M\} \cap \{\omega: \eta(\omega) \in \bigcap_{n=1}^{\infty} Q(\tau_n(\omega))\}).$$

Using the conditions (3), (6) and (7), we see that (8) holds, hence, the first assertion of Theorem 6 is an immediate consequence of (8). Since, in addition,

$$\mu(\{\omega: \ell(\omega)=\infty\} | \{\omega: \eta(\omega) \in A-M\}) =$$
$$= \frac{1}{1-\alpha} \mu(\{\omega: \eta(\omega) \in A-M\} \cap \{\omega: \eta(\omega) \in \bigcap_{n=1}^{\infty} Q(\tau_n(\omega))\},$$

the second assertion of Theorem 6 follows at once from (8).

Let us note that under the assumptions of the theorem just proved it is not true that the conditional moments of ℓ under the condition that the value of η belongs to $A-M$ are finite i.e. the analogue of Theorem 2 in [7] does not hold. This disadvantage, however, can be removed by adding further restrictive conditions.

§3. Semantic concepts.

The statistical decision problem of §2 is based on observations on the sets $Q(1), Q(2), Q(3), \dots$ which replace the direct observation on M. The most natural way of how to get the sets $Q(1), Q(2), Q(3), \dots$ is the effect of reduction of resolving power in A which can be formally described by an appropriate application of equivalence relations.

A binary relation R in the set A is said to be an equivalence relation if it is reflexive, symmetric and transitive. Every equivalence relation in A induces a partition of A into equivalence sets and vice versa. Two elements p, q of A belong to the same equivalence set if and only if pRq. The equivalence relation S in A is said to be finer than R and we shall write $S \leq R$ if pSq implies pRq or, in other words, if every equivalence set induced by S is included in at least one, hence, in exactly one equivalence set induced by R. Clearly, the set of all equivalence relations in A is partially ordered by $\leq$ and the identity I is the finest equivalence relation.

The formal description of reduction of resolving power by equivalence relations is intuitively justified by the convention that two elements of A which belong to the same equivalence set cannot be distinguished. Under this convention it is reasonable to introduce the concept of closure M^R of M induced by the equivalence relation R, requiring that

$$M^R = \bigcup_{q \in M} \{p : pRq\}$$

Clearly, $M^I = M$ i.e. the application of the identity I on M has no effect, and $M^R \subset M^S$, whenever $R \leq S$.

Let $R_1, R_2, R_3, \dots$ be a sequence of equivalence relations in A. Putting

$$M^{R_n} = Q(n)$$

for $n = 1, 2, 3, \dots$, we see that, in fact, the decision problem of §2 is based on observations at a reduced resolving power. This artificial reduction of resolving power is justified by the fact that, in general, M^R has a simpler structure than M^S, whenever $R \geq S$.

The application of the elementary facts summarized in §2 to our main question of statistical estimation of provability by interpretations in models requires a number of restrictions which must be imposed on the sets A and M and on the equivalence relations $R_1, R_2, R_3, \dots$

First of all we shall suppose that A is a Boolean algebra. As usual, we shall denote by 0 and 1 the zero and unity of A, by p' the complement of the element p of A, by $\wedge$ the operation of forming the greatest lower bound, and by $\vee$ that of forming the least upper bound in A.

In order to eliminate misunderstandings we recall that the subset M of A is said to be a Boolean ideal in A if it contains the greatest lower bound of any two of its elements as well as the least upper bound of any two elements of A one at least of which belongs to M. The algebraic structure just defined is usually called dual Boolean ideal. We shall, however, omit the suffix dual because no other ideals will occur in this paper.

The relation R defined in A is said to be a Boolean congruence relation if it is an equivalence relation which, in addition, satisfies the condition

(9) $p \, R \, q$ implies $p' \vee r \, R \, q' \vee r$.

The simplest algebraic structure, which enables the treatment of propositional functions of mathematical logic and for which the concept of interpretation is natural, is that of monadic algebra introduced by P.R.Halmos [2].

A monadic algebra is a Boolean algebra A together with an operator $\forall$ which assigns to every element p of A an element $\forall p$ of A in such a way that

$$\forall 1 = 1,$$
$$\forall p \leq p$$

for every element p of A, and

$$\forall (p \vee \forall q) = \forall p \vee \forall q$$

for every p and q in A. The operator $\forall$ is said to be the universal quantifier in A.

A subset M of a monadic algebra A is said to be a monadic ideal in A whenever M is a Boolean ideal in A and

$$p \in M \text{ implies } \forall p \in M.$$

A monadic ideal M in the monadic algebra A is called maximal if it is proper i.e. $M \neq A$ and it is not a proper subset of any other proper ideal in A.

The relation R defined in the monadic algebra A is said to be a monadic congruence relation if it is a Boolean congruence relation and, in addition,

(10) $p \, R \, q$ implies $\forall p \, R \, \forall q$.

A monadic congruence relation R in the monadic algebra A is called simple whenever the monadic residual class algebra $A(R)$ of A modulo R is simple i.e. there is no proper monadic ideal in A other than that containing the sole element 1.

The relevant properties of closures of monadic ideals will be expressed by the following lemma:

<u>If M is a monadic ideal and R a monadic congruence relation in the monadic algebra A then the closure M^R of M induced by R is a monadic ideal in A. If, in addition, R is simple then either $M^R = A$ or M^R is maximal.</u>

We shall first show that M^R is a Boolean ideal in A. Let $p \in M^R$ and $q \in A$. By the definition of the closure M^R of M, there exists an element r_p of M such that $r_p \, R \, p$. By (9) $r_p' \vee 0 \, R \, p' \vee 0$ i.e. $r_p' \, R \, p'$, hence, $r_p \vee q \, R \, p \vee q$. Since M is a Boolean ideal in A, we have $r_p \vee q \in M$, hence $p \vee q \in M^R$. Now let us suppose that, in addition, $q \in M^R$. Then there exists an element r_q of M such that $r_q \, R \, q$. By (9) we have $r_p' \vee r_q' \, R \, p' \vee r_q'$, $p' \vee r_q' \, R \, p' \vee q'$, hence,

$$(r_p' \vee r_q')' \vee 0 \;\; R \;\; (p' \vee r_q')' \vee 0,$$
$$(p' \vee r_q')' \vee 0 \;\; R \;\; (p' \vee q')' \vee 0$$

or, equivalently,

$$r_p \wedge r_q \;\; R \;\; p \wedge r_q \;,$$
$$p \wedge r_q \;\; R \;\; p \wedge q \;,$$

and, using the transitivity of R, we obtain

$$r_p \wedge r_q \; R \; p \wedge q.$$

Since M is a Boolean ideal in A we have $r_p \wedge r_q \in M$, hence, $p \wedge q \in M^R$. We see that M^R is a Boolean ideal in A. Now it remains to show that M^R is a monadic ideal in A i.e. that $\forall p \in M^R$, whenever $p \in M^R$. Since $r_p \, R \, p$, it follows from (10) that $\forall r_p R \forall p$, hence, using the assumption that M is a monadic ideal in A, we have $\forall r_p \in M$ and, consequently, $\forall p \in M^R$. This completes the proof of the first part of our lemma. If R is simple then, by the definition of simplicity, the class of all congruence sets which have a non-empty intersection with M is either equal to $A(R)$ or to the monadic ideal $\{1\}$ in $A(R)$, hence, either $M^R = A$ or $M^R = \{p : p \, R \, 1\}$, q.e.d.

A monadic logic is a pair (A, M), where A is a monadic algebra and M is a monadic ideal in A. The monadic logic (A, M) represents a deductive theory in A. The elements of A which belong to M are called provable. If R is a simple congruence relation in A then the closure M^R of M induced by R is said to be an interpretation of M in the model $A(R)$. If an element p of A belongs to the interpretation M^R of M we shall say that p is true in that interpretation and otherwise false.

The monadic logic (A, M) is said to be semantically consistent if there exists at least one interpretation of M in a model.

Since $M \subset M^R$, we can state that a provable element of A is true in every interpretation. Whenever the opposite conclusion is possible then the the monadic logic (A, M), is called semantically complete. More precisely, the monadic logic (A, M) is said to be semantically complete if M is equal to the intersection of all its interpretations.

For our purposes, however, a restricted version of semantic completeness is more appropriate. Let $\mathcal{Q}$ be a class of interpretations of M. The monadic logic (A, M) is said to be semantically $\mathcal{Q}$-complete, whenever

$$M = \bigcap_{Q \in \mathcal{Q}} Q .$$

In order to eliminate degenerate cases it is natural to assume that the monadic logic (A, M) is semantically consistent and semantically $\mathcal{Q}$-complete. Clearly, the assumption of semantic consistency can be replaced by $M \neq A$ and, by our lemma, there is no restriction of generality if we assume that the interpretations from $\mathcal{Q}$ are maximal monadic ideals.

The estimation of provability or non-provability of elements of a monadic logic is based upon the inspection of its truth or falsehood in interpretations in models. Since to each interpretation Q from $\mathcal{Q}$ there corresponds a simple monadic congruence relation R_Q such that $Q = M^{R_Q}$, the idea of artificial reduction of resolving power by simple monadic congruence relations is justified by the fact that, by the lemma, the induced closures are maximal monadic ideals which evidently have an extremely simple algebraic structure.

The application of the results established in §2 to the question of statistical estimation of provability in monadic logic requires a further restriction, namely, that $\mathcal{Q}$ is demurable. In this case we can write

$$\mathscr{U} = \{ Q(1),\ Q(2),\ Q(3), \ldots \}.$$

The random variable η chooses an element of the monadic algebra A, the provability or non-provability of which is to be estimated on the basis of interpretations of M chosen from $\mathscr{U}$ by the random variables $\tau_1, \tau_2, \ldots, \tau_n$.

One may intuitively expect that the following decision procedure is the most favorable one. At the first step we choose the interpretation $Q(m)$ determined by the value of τ_1. If the value of η is false in this interpretation, the procedure is stopped and the value of η is estimated to be non-provable. In the opposite case we continue the inspection choosing the interpretation $Q(m)$ determined by the value of τ_2. If the value of η is false in this interpretation, the procedure is stopped and the value of η is estimated to be non-provable. In the opposite case we continue the inspection choosing the interpretation $Q(m)$ determined by τ_3 and so on. Exhausting all the interpretations $Q(m)$ determined successively by the values of $\tau_1, \tau_2, \ldots, \tau_n$ without reaching the decision that the value of η is non-provable we accept the decision that the value of η is provable.

In fact, the decision procedure just described minimizes the probability of making an incorrect decision only if the a priori probability α that η chooses a provable element of A is sufficiently large. Whenever α is small then the degenerate decision procedure which always estimates the value of η to be non-provable is better. The exact discrimination between these two decisions procedures is contained in Theorem 4.

If the a priori probability α is positive, if the values of the random variables $\tau_1, \tau_2, \tau_3, \ldots$ exhaust with probability one whole set of positive integers, and if the monadic logic (A, M) is $\mathscr{U}$-complete then, by Theorem 5, for a sufficiently large number of interpretations to be inspected, the non-degenerate estimation procedure is the most favorable one in the sense that the probability of making an incorrect estimate becomes a minimum. Let us note that the condition of semantic consistency is in this case fulfilled automatically as always whenever $\alpha > 0$ and (A,M) is semantically $\mathscr{U}$-complete.

In the language of monadic logic the decision function β of infinite size occurring in Theorem 6 is said to be the heuristic reasoning about the element of A chosen by η and the random variable ℓ is called the length of the heuristic reasoning β.

The content of Theorem 6 can be expressed as follows: If $\alpha > 0$, if the values of the random variables $\tau_1, \tau_2, \tau_3, \ldots$ exhaust with probability one whole set of positive integers and if the monadic logic (A, M) is semantically $\mathscr{U}$-complete, then the length of the heuristic reasoning about the value of η is infinite with conditional probability one under the condition that a provable element of A has been chosen by η and it is finite with conditional probability one under the condition that the element of A chosen by η was non-provable.

Clearly, only the last assertion is practically effective because only non-provability can be discovered after a finite number of steps. On the other hand, this pessimistic opinion concerning heuristic reasoning is weakened by the fact that if provability is estimated then this result is asymptotically good.

References.

[1] J.L.Doob: Stochastic processes. New York 1953.

[2] P.R.Halmos: Algebraic logic I, Monadic algebras. Comp.Math.12,(1955),pp.217-249.

[3] P.R.Halmos: Algebraic logic II, Homogeneous locally finite polyadic Boolean algebras of infinite degree. Fundam.Math.XLIII,(1950),pp.255-325.

[4] P.R.Halmos: Measure theory. New York 19

[5] J.Neyman and E.S.Pearson: Contributions to the theory of testing statistical hypotheses. Stat. Research Mem., parts I and II (1936 and 1938).

[6] A.Perez: Transformation ou sigma-algèbre suffisante et minimum de la probabilité d'erreur. Czechoslovak Math.J.7,(1957),pp.115-123.

[7] A.Špaček: Statistical estimation of provability in Boolean logic. Transactions of the Second Prague Conference on Information Theory, Statistical Decision Functions and Random Processes, 1960,pp.

ON RANDOM OPERATOR EQUATIONS

Otto Hanš

25X1

Approved For Release 2009/07/09 : CIA-RDP80T00246A011700340001-4

Introduction.

In the present paper we deal with random operator equations, in particular with random integral equations of Fredholm and Volterra type, where only the kernel is random. It can be shown that such a model is general enough to cover other cases, for example random limits of integration.

However, the author was not able to give a more detailed discussion of the subject, mainly because of the lack of time. Nevertheless, he prefers this somewhere even blurring preliminary version to a short summary.

Random operator equations.

First of all we shall introduce some basic conceptions indispensable for our further considerations.

Throughout the whole paper we shall use a probability space $(\Omega, \mathfrak{S}, \mu)$ with a complete probability measure μ, i.e. Ω is a non-empty set, $\mathfrak{S}$ a σ-algebra of subsets of the set Ω, and μ is a probability measure such that $\mu(E_0) = 0$ implies $E \in \mathfrak{S}$ for every $E \subset E_0$. Further, unless there will be a statement to the contrary, X denotes an arbitrary separable Banach space and $\mathfrak{X}$ the σ-algebra of all Borel subsets of the space X.

D e f i n i t i o n 1. A mapping V of the space Ω into the space X is called a generalized random variable if

$$\{\{\omega : V(\omega) \in B\} : B \in \mathfrak{X}\} \subset \mathfrak{S}.$$

D e f i n i t i o n 2. A mapping T of the Cartesian product $\Omega \times Z$ into the space X is called a random transformation if $T(\cdot, z)$ is for every $z \in Z$ a generalized random variable.

Two generalized random variables V and W are assumed to be equivalent if

$$\mu\{\omega : V(\omega) = W(\omega)\} = 1.$$

All the classical notions of the functional analysis, like the inverse operator, the resolvent operator, the adjoint operator, etc., are carried over in an "almost sure" way, that is, e.g., the mapping S is said to be the inverse of the mapping T if

$$\mu\{\omega : S(\omega, T(\omega, z)) = z \text{ for every } z \in Z\} = 1.$$

Finally, we recall the following three theorems on generalized random variables :

T h e o r e m 1. If $V_1, V_2, \ldots$ is a sequence of generalized random variables converging almost surely to the mapping V, then V is a generalized random variable.

T h e o r e m 2. A mapping V of the space Ω into the space X is a generalised random variable if, and only if, for every bounded linear functional $f \in \Delta$, where Δ is a subset of the first adjoint space X^* which is total on the whole space X, the compound mapping $f(V)$ is a real-valued random variable.

T h e o r e m 3. Let V be a generalized random variable with values in the space X and T a random transformation of the Cartesian product $\Omega \times X$ into the space Z that is almost surely continuous. Then the mapping W of the space Ω into the space Z defined for every $\omega \in \Omega$ by the formula

$$W(\omega) = T(\omega, V(\omega))$$

is a generalized random variable with values in the space Z.

Let X and Z be two separable Banach spaces, T a mapping of the space X into the space Z, and z a fixed element from Z. If Σ denotes the set of those elements $x \in X$ for which the equality $T(x) = z$ holds, i.e. if

$$\Sigma = \{x : T(x) = z\}$$

then any $x \in \Sigma$ is used to be called a solution of the operator equation

$$T(\xi) = z . \qquad /1/$$

If the set Σ is empty, we say that the operator equation /1/ does not possess a solution; if it is non-empty, we say that /1/ is solvable. In the case Σ consists of exactly one point we say that /1/ has a unique solution.

Now, let in addition $(\Omega, \mathfrak{S}, \mu)$ be a probability space with a complete probability measure μ and $\mathfrak{X}$ and $\mathfrak{Z}$ be the σ-algebras of all Borel subsets of the space X and Z respectively. If T is a random transformation of the Cartesian product $\Omega \times X$ into the space Z then

$$T(\cdot, \xi) = z \qquad /2/$$

is said to be random operator equation.

However, the relation /2/ does not express the most general form of a random operator equation, namely the right hand side of /2/ need not be a fixed element from the space Z, but can be replaced by a generalized random variable with values in the space Moreover, it should be remarked that the solution of a random operator equation does in general depend on the choice of $\omega \in \Omega$; consequently, the most general form of the random operator equation may be more precisely written as

$$T(\cdot, \xi(\cdot)) = z(\cdot) \qquad /3/$$

where, as mentioned above, T is a random transformation of the Cartesian product $\Omega \times X$ into the space Z and z is a generalized random variable with values in the space Z.

Similarly as in the deterministic case, with the only exception of neglecting a set of probability measure zero, the wide sense solutions of the equation /3/ are defined, i.e. every mapping w of the space Ω into the space X satisfying the equality

$$T(\omega, w(\omega)) = z(\omega)$$

for every ω from a set Ω_0 of probability measure one is said to be a wide sense solution of the equation /3/. However, following the spirit of our previous papers, it is quite natural to require the condition of measurability to be fulfilled in order that we can speak about random solutions. Thus, if the wide sense solution is moreover measurable it will be called the random solution, and we can state

D e f i n i t i o n 3. Every generalized random variable x with values in the space X satisfying the condition

$$\mu\{\omega : T(\omega, x(\omega)) = z(\omega)\} = 1$$

will be called the random solution of the random operator equation /3/.

Evidently, there may exist wide sense solutions that are not random solutions. Moreover, if the random operator equation has more than one solution for every ω from a set of positive probability measure then they may be, in dependance on the σ-algebra $\mathfrak{S}$

of course, many wide sense solutions that are not measurable. As a simple example of this fact let us mention

Example 1. Let X be the space of all real numbers, E a non-measurable subset of the space Ω and T a random transformation of the Cartesian product $\Omega \times X$ into the space X, defined for every $\omega \in \Omega$ and $x \in X$ by the formula

$$T(\omega, x) = x^2 - 1 .$$

Then the mapping v of the space Ω into the space X defined by the formulae

$$v(\omega) = 1 \qquad \text{for } \omega \in E ;$$

$$v(\omega) = -1 \qquad \text{otherwise} ;$$

is a wide sense solution, but is not a random solution, of the random operator equation

$$T(\cdot, \xi(\cdot)) = 0 . \qquad /4/$$

Roughly speaking, we are therefore interested mainly in the case when for every $\omega \in \Omega$ there exists a unique solution of the deterministic operator equation

$$T(\omega, \xi) = z(\omega) .$$

More precisely, we shall investigate most frequently the case when here exists a unique wide sense solution, provided we identify t...

solutions differring on a set of probability measure

Nevertheless, even under this restriction, the uni

se solution need not be measurable as shown by the following

Example 2. Let Ω be the space of all real numbers, $\mathfrak{S}$ the σ-algebra of all at most denumerable sets of real numbers and their complements, and μ a complete probability measure defined by the formulae

$\mu(E) = 0$ if E is at most denumerable;

$\mu(E) = 1$ if the complement of E is at most denumerable:

Further, let X be the space of all real numbers with the σ-algebra $\mathcal{X}$ of all Borel subsets, and T a random transformation of the Cartesian product $\Omega \times X$ into the space X defined by the formulae

$T(\omega, x) = 0$ for $\omega = x$;

$T(\omega, x) = 1$ otherwise.

Then the unique wide sense solution of the random operator equation /4/, given for every $\omega \in \Omega$ by the formula

$$v(\omega) = \omega$$

is not a random solution, because of the fact that, e.g.,

$$\{\omega : v(\omega) \leq 0\} \notin \mathfrak{S}.$$

Many other questions concerning the relationship between wide sense solutions and random solutions of the same random operator equations arise, the greater part of them having been as yet unsolved. Unfortunately, we too are not able to present some useful theory of random operator equations unless some further assumptions are imposed.

In the present paper we shall discuss mainly three particular cases, namely if

a/ the separable Banach space X equals to the separable Banach space Z ; and/or if

b/ the random transformation T is almost surely linear and bounded.

It is not surprising that in the theory of random operator equations the main role is played by probabilistic versions of the well known Principle of Contraction Mappings and its many modifications and generalizations, which under appropriate assumptions furnish the existence, uniqueness, and measurability of the random solution of a random operator equation.

The following theorem is a useful starting point for other theorems of this kind.

T h e o r e m 4. Let T be an almost surely continuous random transformation of the Cartesian product $\Omega \times X$ into the space X so that

$$\mu\Big(\bigcup_{m=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{x\in X}\bigcap_{y\in X}\{\omega: \|T^{n}(\omega,x)-T^{n}(\omega,y)\| \le (1-\tfrac{1}{m})\cdot\|x-y\|\}\Big)=1, \qquad /5/$$

where for every $\omega \in \Omega$, $x \in X$, and $n = 1, 2, \cdots$

$$T^{1}(\omega,x) = T(\omega,x)$$

and

$$T^{n+1}(\omega,x) = T(\omega, T^{n}(\omega,x)).$$

Then there exists a generalized random variable ϕ with values in the space X satisfying the relation

$$\mu\{\omega: T(\omega,\phi(\omega)) = \phi(\omega)\} = 1. \qquad /6/$$

Moreover, if there exists another generalised random variable ψ with the property

$$\mu\{\omega: T(\omega,\psi(\omega)) = \psi(\omega)\} = 1$$

then

$$\mu\{\omega: \phi(\omega) = \psi(\omega)\} = 1.$$

P r o o f. Let us denote by E the set of those ω from the set occurring in condition /5/ for which the mapping $T(\omega,\cdot)$ is continuous. Evidently, according to the assumptions, $\mu(E) = 1$. Now, let us define the mapping ϕ of the space Ω into the space X so that for every $\omega \in E$ the point $\phi(\omega)$ equals to the unique fixed point of the mapping $T(\omega,\cdot)$ and for every $\omega \in \Omega - E$ we set $\phi(\omega) = \theta$, where θ is the null element of the Banach space X. Thus, relation /6/ holds.

In order to prove the measurability of the mapping ϕ we make use of theorem 3 and theorem 1. The remaining statement follows immediately from the uniqueness of the fixed point of the mapping $T(\omega,\cdot)$ for every $\omega \in E$.

The just proved theorem forms a generalisation of author's previous result, the stronger assumptions of which enable one to formulate condition /5/ in a more transparent way.

T h e o r e m 5. Let T be an almost surely continuous random transformation of the Cartesian product $\Omega \times X$ into the space X and c a real-valued random variable so that the following relations hold :

$$\mu\{\omega : c(\omega) < 1\} = 1 ; \qquad /7/$$

$$\left.\begin{array}{l}\mu\{\omega : \|T(\omega,x) - T(\omega,y)\| \leq c(\omega) \cdot \|x-y\|\} = 1 \\ \text{for every two elements } x \text{ and } y \text{ from } X .\end{array}\right\} \qquad /8/$$

Then there exists a generalized random variable ϕ with values in the space X for which relation /6/ holds.

P r o o f. Since the Banach space X is separable, we can replace the intersections in the expression

$$\bigcap_{x \in X} \bigcap_{y \in X} \left[\{\omega : \|T(\omega,x) - T(\omega,y)\| \leq c(\omega)\|x-y\|\} \cap \{\omega : c(\omega) < 1\} \cap \{\omega : T(\omega,\cdot) \text{ is continuous}\}\right]$$

by intersections over a countable dense subset of the space X and thus condition /5/ is fulfilled with $n = 1$, what proves our theorem.

R e m a r k 1. It should be remarked that both theorem 4 and theorem 5 can be reformulated and proved also in the case the space X under consideration is a metric space. However, in the present paper we formulate all the theorems for Banach spaces only, though the "metric" versions of theorems 4 and 5 are used in proofs of theorems 11 and 12 in order to avoid more complicated considerations.

As an immediate consequence of theorem 5 we get

T h e o r e m 6. Let C be a real-valued random variable and T an almost surely random transformation of the Cartesian product $\Omega \times X$ into the space X satisfying the condition /8/.

Then for every real number $\lambda \neq 0$ such that

$$\mu\{\omega : C(\omega) < |\lambda|\} = 1$$

there exists a random transformation S that is the inverse of the random transformation $(T - \lambda I)$.

P r o o f. Evidently, as $\lambda \neq 0$, the random transformation $(T - \lambda I)$ is invertible whenever the random transformation $(\frac{1}{\lambda} T - I)$ is invertible. However, for every $z \in X$ the random transformation T_z defined for every $\omega \in \Omega$ and $x \in X$ by the formula

$$T_z(\omega, x) = \frac{1}{\lambda} T(\omega, x) - z$$

is almost surely reducing and therefore by theorem 5 there exists a unique random fixed point x_z satisfying the relation

$$\mu\{\omega : x_z(\omega) = \frac{1}{\lambda} T(\omega, x_z(\omega)) - z\} = 1.$$

Since the last statement is equivalent to the invertibility of the random transformation $(\frac{1}{\lambda} T - I)$, theorem 6 is proved.

It is a well known fact that any linear bounded operator A satisfies the Lipschitz condition with the constant $\|A\|$, what is at the same time the smallest constant with such a property. Therefore making use of this fact and some classical results about linear bounded operators we can state

T h e o r e m 7. Let T be an almost surely linear bounded random transformation of the Cartesian product $\Omega \times X$ into the space X .

Then for every real number $\lambda \neq 0$ such that

$$\mu\left(\bigcup_{n=1}^{\infty}\{\omega: \|T^n(\omega,\cdot)\| < |\lambda|^n\}\right) = 1$$

there exists a linear bounded random transformation S that is the inverse of the random transformation $(T - \lambda I)$ and we have

$$\mu\left(\bigcap_{x\in X}\{\omega: S(\omega,x) = -\frac{1}{\lambda}\sum_{n=0}^{\infty}\frac{1}{\lambda^n}T^n(\omega,x)\}\right) = 1,$$

where the sum is meant uniformly.

A number of interesting theorems on random integral equations has been derived by A.T. Bharucha-Reid. Here we mention theorems 2.1 through 2.3 only, which can be a little strenghtened using theorem 7, namely we can write

T h e o r e m 8. Let T be a random transformation of the Cartesian product $\Omega \times X$ into the space X which is for every $\omega \in \Omega$ linear and bounded.

Then for every real number $\lambda \neq 0$ the set

$$\Omega(\lambda) = \{\omega: \|T(\omega,\cdot)\| < |\lambda|\}$$

belongs to the σ-algebra $\mathfrak{S}$, the random transformation $(T-\lambda I)$ is invertible for every $\omega \in \Omega(\lambda)$; the resolvent operator $R_{\lambda,T}$ exists for every $\omega \in \Omega(\lambda)$ and for these ω's we have

$$R_{\lambda,T}(\omega,\cdot) = -\sum_{n=1}^{\infty}\lambda^{-n}\,T^{n-1}(\omega,\cdot).$$

Finally, for every $\omega \in \Omega(\lambda)$ the solution $s(\omega)$ of the operator equation

$$T(\omega,\xi) - \lambda\xi = z$$

is for every $z \in X$ given by the formula

$$s(\omega) = R_{\lambda,T}(\omega,z) \, ,$$

where the resolvent operator $R_{\lambda,T}$, and consequently the solution s as well, is measurable with respect to the σ-algebra $\Omega(\lambda) \cap \mathfrak{S}$.

P r o o f. Theorem 8 follows immediately from our preceding theorems and from well known classical results, because of the fact that for every $\omega \in \Omega$

$$\| T(\omega,\cdot) \| = \sup_{n \geq 1} \| T(\omega, x_n) \|$$

where $x_1, x_2, \ldots$ is a countable dense subset of the sphere $\{x : \|x\| \leq 1\}$.

Heretofore we have always used the above formulated assumption a/, sometimes together with the assumption b/. Now, we shall state two theorems in which the spaces X and Z may be different separable Banach spaces, provided the random transformation T is almost surely linear and bounded. The proofs are omitted, because both the theorems are only slightly modificated previous author's results.

T h e o r e m 9. The inverse of an almost surely linear bounded invertible random transformation T of the Cartesian product $\Omega \times X$ into the space Z is a random transformation of the Cartesian product $\Omega \times Z$ into the space X .

T h e o r e m 10. Let X and Z be two Banach spaces whose first adjoint spaces X^* and Z^* are separable; let $\mathfrak{X}$, $\mathfrak{Z}$, $\mathfrak{X}^*$, and $\mathfrak{Z}^*$ be the σ-algebras of all Borel subsets of the space X , Z , X^*, and Z^* respectively. If T is an almost surely linear bounded mapping of the Cartesian product $\Omega \times X$ dnto the space Z then the following two conditions are equivalent :

for almost all elements $\omega \in \Omega$ the mapping $T(\omega,\cdot)$ of the space X onto the space Z is invertible ;

for almost all elements $\omega \in \Omega$ the range of the adjoint mapping $T^*(\omega,\cdot)$ is the whole space X^* .

Further, if these conditions are satisfied then T^* is invertible and the inverse mapping $(T^*)^{-1}$ to the adjoint mapping T^* is almost surely equal to the adjoint mapping S^* of the inverse S to the mapping T .

Moreover, if one of the mappings T , S , T^*, S^* is a random transformation then all four mappings are random transformations.

One of the important problems in the theory of random operator equations is the question of the measurability of the solution, which has been dealt with in the preceding theorems. Now, we shall be concerned with the question of the relationship between the random solution of the random operator equation and the solution of the corresponding deterministic operator equation. More precisely, we shall discuss the case when the random operator equation /3/ is such that the Bochner integrals

$$\int_\Omega z(\omega)\, d\mu(\omega) = z$$

and

$$\int_\Omega T(\omega, x)\, d\mu(\omega) = S(x)$$

exist for every $x \in X$. Let us assume that the solution of the deterministic operator equation

$$S(\xi) = z$$

equals to y . The question arises whether the expected value of the random solution of the random operator equation /3/ exists, and if so, whether it is equal to the deterministic solution y .

It is not difficult to construct an example showing that there are cases in which the answer is affirmative. A most trivial one is that when the probability measure μ is a Dirac measure, i.e. when there exists an element $\omega_0 \in \Omega$ such that $\mu(\omega_0) = 1$. Another still trivial example is the following

E x a m p l e 3. Let T be a random transformation of the Cartesian product $\Omega \times X$ into the space X defined for every $\omega \in \Omega$ and $x \in X$ by the formula

$$T(\omega, x) = cx + V(\omega),$$

where $c \neq 0$ is a real number and V a generalized random variable with values in the space X so that the Bochner integral

$$\int_\Omega V(\omega)\, d\mu(\omega)$$

equals to the null element of the space X.

Then the expected value of the unique random solution of the random operator equation

$$T(\cdot, \xi(\cdot)) = \theta$$

equals to the solution of the operator equation

$$S(\xi) = \theta$$

where, of course,

$$S(x) = cx$$

for every $x \in X$.

It is also not difficult to give examples when the answer to the above stated question is negative.

However, there are many cases in which we are interested in the solution of the deterministic operator equation corresponding to a given random operator equation rather than in the expected value, or even in the probability distribution, of the random solution of the random operator equation under consideration. Really, the case of determining LD 50 is one of them.

Let us call, for the sake of brevity, the deterministic operator equation associated with the random operator equation by means of expected-value correspondance simply the regression operator equation.

Thus, an important problem of the theory of random operator equations is that of reaching the solution of the regression operator equation, and it is this problem the remainder of this section is devoted to.

Since a detailed case history of this and similar problems is given in a common paper written by M. Driml and the author which will appear in the Transactions of the Second Prague Conference on Information Theory, Statistical Decision Functions, and Random Processes, we shall not go into details here, but shall state three useful theorems only. Nevertheless, it should be noted that this branch of probability theory, often called experience theory, is very tightly connected with stochastic approximation methods as developed by H. Robbins and S. Monro and other authors.

The following theorem, as yet unpublished, is due to M. Driml and the author and can be counted to the basic theorems of the experience theory.

T h e o r e m 11. Let T be a generalized stochastic process mapping the Cartesian product $\Omega \times [0,\infty) \times X$ into the space X and almost surely continuous with respect to both the arguments $t \geq 0$ and $x \in X$ simultaneously. Let there exist an element $\hat{x} \in X$, a real-valued random variable c, and let the following conditions, together with condition /7/, hold :

$$\mu\{\omega : \lim_{t\to\infty} \| \tfrac{1}{t}\int_0^t T(\omega, s, \hat{x})\, ds - \hat{x} \| = 0\} = 1; \qquad /9/$$

$$\mu\{\omega : \|T(\omega,t,x) - T(\omega,t,y)\| \le c(\omega) . \|x-y\|\} = 1$$

for every $t \geq 0$ and every two elements $x, y \in X$.

Further, let $x_t(\cdot)$ be the solution of the random operator equation

$$\left.\begin{aligned} \xi_0(\cdot) &= T(\cdot, 0, \xi_0(\cdot)) \\ \xi_t(\cdot) &= \frac{1}{t}\int_0^t T(\cdot, s, \xi_s(\cdot))\, ds \quad \text{for } t > 0 . \end{aligned}\right\} /10/$$

Then x_t is for every $t \geq 0$ a random solution of the random operator equation /10/ and we have

$$\mu\{\omega : x_t(\omega) \text{ is continuous in } t\} = 1 \qquad /11/$$

and

$$\mu\{\omega : \lim_{t\to\infty} \| x_t(\omega) - \hat{x} \| = 0\} = 1. \qquad /12/$$

P r o o f. Let us denote by C the space of all continuous mappings f of the space $[0,\infty)$ into the space X such that the relation

$$\lim_{t\to\infty} \|f(t) - \hat{x}\| = 0$$

holds. Introducing the distance function δ for every couple of elements $f, g \in C$ by the formula

$$\delta(f,g) = \sup_{t \geq 0} \|f(t) - g(t)\|$$

the space C becomes a separable metric space whose σ-algebra of all Borel subsets is the σ-algebra generated by the class

$$\{\{f : f(t) \in B\} : 0 \le t < \infty, B \in \mathfrak{X}\}.$$

Further, let us denote by S the operator on the Cartesian product $\Omega \times C$ defined by the formulae

$$[S(\omega,f)](0) = T(\omega,0,f(0))$$

and

$$[S(\omega,f)](t) = \frac{1}{t}\int_0^t T(\omega,s,f(s))\,ds$$

for every $f \in C$, $t > 0$, and every $\omega \in E$, where E equals to the intersection

$$\{\omega : T(\omega,t,x) \text{ is continuous in } t \text{ and } x \text{ simultaneously}\} \cap$$
$$\cap \{\omega : c(\omega) < 1\} \cap \{\omega : \lim_{t\to\infty} \| \frac{1}{t}\int_0^t T(\omega,s,\hat{x})\,ds - \hat{x} \| = 0\} \cap$$
$$\cap \bigcap_{t\geq 0} \bigcap_{x\in X} \bigcap_{y \in X} \{\omega : \|T(\omega,t,x) - T(\omega,t,y)\| \leq c(\omega).\|x-y\|\} .$$

For every $\omega \in \Omega - E$, $f \in C$, and $t \geq 0$ let us put

$$[S(\omega,f)](t) = \hat{x} .$$

First of all let us prove that the mapping S maps $\Omega \times C$ into C . Choose arbitrarily $\omega \in E$, $f \in C$, and $\varepsilon > 0$. Then there exists a real number t_0 such that for every $t \geq t_0$

$$\| f(t) - \hat{x} \| \leq \varepsilon/3$$

and simultaneously

$$\| \frac{1}{t}\int_0^t T(\omega,s,\hat{x})\,ds - \hat{x} \| \leq \varepsilon/3 .$$

Then for every

$$t \geq t_0 . \delta(f,\hat{x}) . 3/\varepsilon$$

/13.

we can write

$$\| \frac{1}{t}\int_0^t T(\omega,s,f(s))\,ds - \hat{x} \| \leq$$
$$\leq \frac{1}{t}\int_0^t \|T(\omega,s,f(s)) - T(\omega,s,\hat{x})\|\,ds + \| \frac{1}{t}\int_0^t T(\omega,s,\hat{x})\,ds - \hat{x} \| \leq$$
$$\leq c(\omega) . \frac{1}{t}\int_0^t \|f(s) - \hat{x}\|\,ds + \varepsilon/3 \leq$$
$$\leq c(\omega) . \frac{t_0}{t} . \delta(f,\hat{x}) + c(\omega) . \frac{t-t_0}{t} . \sup_{t_0 \leq s \leq t} \|f(s) - \hat{x}\| + \varepsilon/3 \leq \varepsilon .$$

Since the $(\Omega - E)$ part is trivial we have proved that S maps the Cartesian product $\Omega \times C$ into the space C.

Now, let us prove that S is a reducing transformation. Thus, let us have $f, g \in C$ and $\omega \in E$ / for $\omega \in \Omega - E$ we get a singular case/. Then

$$\delta(S(\omega,f), S(\omega,g)) =$$
$$= \sup_{t \geq 0} \left\| \frac{1}{t}\int_0^t T(\omega,s,f(s))\,ds - \frac{1}{t}\int_0^t T(\omega,s,g(s))\,ds \right\| \leq$$
$$\leq \sup_{t \geq 0} \frac{1}{t}\int_0^t \sup_{s \geq 0} \| T(\omega,s,f(s)) - T(\omega,s,g(s)) \|\,ds \leq$$
$$\leq c(\omega) \cdot \delta(f,g).$$

The mapping S being a reducing random transformation of the Cartesian product $\Omega \times C$ into the space C we can, making use of Remark 1, apply theorem 5 which asserts that there exists a generalized random variable with values in the space C so that

$$\mu\{\omega : S(\omega, \psi(\omega)) = \psi(\omega)\} = 1.$$

However, we have $E \in \mathfrak{S}$, $\mu(E) = 1$; hence setting for every $\omega \in \Omega$ and every $t \geq 0$

$$x_t(\omega) = [\psi(\omega)](t)$$

we get immediately all the assertions of theorem 11.

A generalization of the preceding theorem for almost surely linear bounded random transformations is the following

T h e o r e m 12. Let T be a generalized stochastic process mapping the Cartesian product $\Omega \times [0,\infty) \times X$ into the space X and almost surely continuous with respect to both the arguments $t \geq 0$ and $x \in X$ simultaneously. Denote by $T_{t_1,\dots,t_n}$ the mapping of the Cartesian product $\Omega \times X$ into the space X defined for every $\omega \in \Omega$, $x \in X$, and $j = 1, 2, \dots$ by the formulae

$$T_{t_1}(\omega, x) = T(\omega, t_1, x)$$

and

$$T_{t_1,\dots,t_{j+1}}(\omega, x) = T(\omega, t_{j+1}, T_{t_1,\dots,t_j}(\omega, x)).$$

Let there exist an element $\hat{x} \in X$ so that condition /9/ and the following conditions are satisfied :

$$\mu\{\omega : T(\omega, t, \alpha x + \beta y) = \alpha T(\omega, t, x) + \beta . T(\omega, t, y)\} = 1$$

for every $t \geq 0$, $x, y \in X$, and any real numbers α and β ;

$$\mu\{\omega : \sup_{t \geq 0} \| T(\omega, t, \cdot) \| < \infty\} = 1;$$

$$\mu\Big(\bigcup_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{x \in X} \bigcap_{y \in X} \bigcap_{t_1, t_2, \dots, t_n \geq 0} \{\omega : \| T_{t_1,\dots,t_m}(\omega, x-y) \| \leq (1 - \tfrac{1}{m}) . \| x - y \| \}\Big) = 1.$$

Further, let $x_t(\cdot)$ be the solution of the random operator equation /10/.

Then x_t is for every $t \geq 0$ the unique random solution of the random operator equation /10/ and the relations /11/ and /12/ hold.

P r o o f. The proof follows that one of theorem 11. We must only replace the inequality /13/ by the inequality

$$t \geq 3 . t_0 \; \delta(f, \hat{x}) . \sup_{t \geq 0} \| T(\omega, t, \cdot) \| \cdot \frac{1}{\varepsilon}$$

and prove the contraction property in the following way :
Denote by S^n the mapping of the Cartesian product $\Omega \times C$ into the space C defined for every $\omega \in \Omega$, $f \in C$, and $n = 1, 2, \dots$ by the formulae

$$S^1(\omega, f) = S(\omega, f)$$

and

$$S^{n+1}(\omega, f) = S(\omega, S^n(\omega, f)) .$$

Then we can for every $f \in C$, every $n = 1, 2, \dots$, every $t > 0$, and almost all ω's write

$$[S^n(\omega, f)](t) = [S(\omega, S^{n-1}(\omega, f))](t) =$$

$$= \frac{1}{t}\int_0^t T(\omega, s_1, [S^{n-1}(\omega,f)](s_1))\, ds_1 =$$

$$= \frac{1}{t}\int_0^t T(\omega, s_1, \frac{1}{s_1}\int_0^{s_1} T(\omega, s_2, [S^{n-2}(\omega,f)](s_2))\, ds_2)\, ds_1 =$$

$$= \frac{1}{t}\int_0^t \frac{1}{s_1}\int_0^{s_1} T(\omega, s_1, T(\omega, s_2, [S^{n-2}(\omega,f)](s_2)))\, ds_2\, ds_1 =$$

$$= \frac{1}{t}\int_0^t \frac{1}{s_1}\int_0^{s_1} T_{s_1,s_2}(\omega, [S^{n-2}(\omega,f)](s_2))\, ds_2\, ds_1$$

and hence by induction

$$[S^n(\omega,f)](t) =$$

$$= \frac{1}{t}\int_0^t \frac{1}{s_1}\int_0^{s_1} \cdots \frac{1}{s_{n-1}}\int_0^{s_{n-1}} T_{s_1,s_2,\dots,s_n}(\omega, f(s_n))\, ds_n \dots ds_2\, ds_1 .$$

Thus, for every ω from a set of probability measure one there exist positive integers m and n such that for any two elements $f, g \in C$ we have

$$\delta(S^n(\omega,f), S^n(\omega,g)) =$$

$$= \sup_{t\geq 0} \left\| \frac{1}{t}\int_0^t \frac{1}{s_1}\int_0^{s_1} \cdots \frac{1}{s_{n-1}}\int_0^{s_{n-1}} T_{s_1,\dots,s_n}(\omega, f(s_n)-g(s_n))\, ds_n \dots ds_1 \right\| \leq$$

$$\leq \left(1-\frac{1}{m}\right)\cdot \delta(f,g)$$

and hence applying the "metric" version of theorem 4 we get the desired result.

Theorems 11 and 12 and many other theorems of the same type claim the fact that under appropriate assumptions the "decision process" converges to the searched fixed point of the regression transformation with probability one. However, in practical situations the statistician scarcely knows whether all the necessary conditions are satisfied. A kind of justification of his decision about the fixed point of the regression transformation is contained in the following

T h e o r e m 13. Let T be a generalized stochastic process of the Cartesian product $\Omega \times [0,\infty) \times X$ into the space X and $\hat{x}$ a fixed element of the space X. Let $x_t(\cdot)$ be the solution of the random operator equation /10/ and let the following relations together with relation /12/ hold :

$$\mu\{\omega : \lim_{t\to\infty} \| T(\omega, t, x_t(\omega)) - T(\omega, t, \hat{x}) \| = 0\} = 1 ;$$

$$\mu\{\omega : \int_0^t \| T(\omega, s, x_s(\omega)) - T(\omega, s, \hat{x}) \| . ds < \infty \} = 1$$

for every $t \geq 0$.

Then the relation /9/ holds.

P r o o f. According to the condition /12/ it suffices to prove that

$$\mu\{\omega : \lim_{t\to\infty} \| \frac{1}{t}\int_0^t T(\omega, s, \hat{x})\, ds - x_t(\omega) \| = 0\} = 1$$

or

$$\mu\{\omega : \lim_{t\to\infty} \| \frac{1}{t}\int_0^t T(\omega, s, \hat{x})\, ds - \frac{1}{t}\int_0^t T(\omega, s, x_s(\omega))\, ds \| = 0\} = 1.$$

However, the last relation follows immediately from the trivial inequality

$$\| \frac{1}{t}\int_0^t T(\omega, s, \hat{x})\, ds - \frac{1}{t}\int_0^t T(\omega, s, x_s(\omega))\, ds \| \leq$$
$$\leq \frac{1}{t}\int_0^t \| T(\omega, s, \hat{x}) - T(\omega, s, x_s(\omega)) \|\, ds$$

which holds with probability one.

Some other theorems dealing with experience theory problems can be found in some papers published in the Transactions of both the First and the Second Prague Conferences.

Random integral equations.

In the remainder of this preliminary version we shall briefly deal with a special case of random linear operator equations, namely with random integral equations.

In this section X denotes the separable Banach space of all continuous functions defined on a closed interval $[a, b]$, where $0 < a < b$, with the norm

$$\|x\| = \max_{a \le u \le b} |x(u)| .$$

Further, let us denote by Δ the set

$$\{(u,v) : a \le u \le b, a \le v \le b\} .$$

First of all we shall recall the well known result from functional analysis, namely

T h e o r e m 14. If $k(u,v)$ is bounded for every $(u,v) \in \Delta$ and if all discontinuity points of k are situated on a finite number of curves

$$v = \varphi_j(u), \qquad j = 1,2,\dots,n$$

where the functions φ_j are continuous, then the formula

$$y(u) = \int_a^b k(u,v)\, x(v)\, dv$$

defines a compact linear operator on the space X into itself.

For the sake of simplicity we shall not work with the most general form of the kernel k, but shall assume only kernels all discontinuity points of which are situated on the curve $v = u$. It would be still simpler to discuss continuous kernels only, however, this would exclude the important class of integral equations of Volterra type. Therefore we have accepted the above stated compromise.

Thus, let us denote by K the space of all kernels k all discontinuity points of which are situated on the curve $v = u$.

Introducing the norm given by the formula

$$\|k\| = \sup |k(u,v)|,$$

where the least upper bound is taken over the set

$$\Delta - \{(u,v) : u = v\},$$

the space K becomes a separable Banach space.

Now, we can state the following two results concerned with the measurability of the mapping T occurring in the theory of random integral equations.

T h e o r e m 15. Let the mapping T of the Cartesian product $\Omega \times X$ into the space X be defined for every $\omega \in \Omega$ and $x \in X$ by the formula

$$T(\omega, x) = \int_a^b k(\omega, \cdot, v) x(v)\, dv, \qquad /14/$$

where k is a mapping of the space Ω into the space K such that

$$\{\omega : k(\omega, u, v) < r\} \in \mathfrak{S}$$

for every $(u,v) \in \Delta$, and every real number r.

Then the mapping T is a compact linear random transformation.

P r o o f. That the mapping /14/ is for every $\omega \in \Omega$ linear and compact follows immediately from theorem 14. Let us therefore prove that it is a random transformation.

Thus, choose an arbitrary $x \in X$. According to theorem 2 it suffices to prove the measurability for every $x \in X$ and every linear bounded functional g_u defined for every $u \in [a,b]$ and $y \in X$ by the formula

$$g_u(y) = y(u),$$

that is, it suffices to prove the measurability of the mapping

$$g_u[T(\omega, x)] = \int_a^b k(\omega, u, v) x(v)\, dv$$

for every fixed $u \in [a, b]$ and $x \in X$. However, by the same theorem 2 we get that the statement

$$k(\cdot, u, v)\, x(v)$$

is a real-valued random variable for every $(u,v) \in \Delta$ and $x \in X$ implies the statement

$$k(\cdot, u, \cdot)\, x(\cdot)$$

is generalized random variable with values in the space X_*. Now, the mapping associating to every $\omega \in \Omega$, every kernel k, and every $x \in X$ the real number

$$\int_a^b k(\omega, u, v)\, x(v)\, dv$$

is a linear bounded functional, say h, on the space X, so that we can write

$$g_u[T(\omega, x)] = h\,[k(\omega, u, \cdot)\, x(\cdot)]$$

and thus according to theorem 2 T is a random transformation.

T h e o r e m 16. The compact linear random transformation T defined in theorem 15 satisfies the relation

$$\{\omega : T(\omega, \cdot) \in B\} \in \mathfrak{S}$$

for every Borel /with respect to the normed topology/ subset B of the space of all endomorphisms on X into itself.

P r o o f. Heretofore we have dealt with the random transformation T as with the mapping of the Cartesian product $\Omega \times X$ into the space X, only. Now, we shall use another point of view.

Let us denote by M the space of all endomorphisms on X into itself, i.e. the space of all linear bounded operations A on X into itself. Introducing the usual norm into this space by the formula

$$\|A\| = \sup_{\|x\|=1} \|A(x)\|$$

the space M becomes a Banach algebra. Denote further by $\mathfrak{M}$ the σ-algebra of all Borel /with respect to the normed topology just introduced/ subsets of the space M. Thus, we have to prove the relation

$$\{\{\omega : T(\omega,\cdot) \in B\} : B \in \mathfrak{M}\} \subset \mathfrak{S}.$$

However, this relation is equivalent to the relation

$$\{\{\omega : T(\omega,\cdot) \in B\} : B \in \mathfrak{F}\} \subset \mathfrak{S},$$

where the σ-algebra $\mathfrak{F}$ is defined by the formula

$$\mathfrak{F} = \{B \cap F : B \in \mathfrak{M}\}$$

and F is the space of all possible compact linear operators formed by means of kernels from the space K, i.e. in symbols

$$F = \{L : L(x) = \int_a^b k(\cdot,v)x(v)\,dv,\ x \in X,\ k \in K\}$$

Because of the inequality

$$\|L\| = \sup_{\|x\|=1} \|L(x)\| =$$

$$= \sup_{\|x\|=1} \|\int_a^b k(\cdot,v)x(v)\,dv\| =$$

$$= \max_{a \le u \le b} \sup_{\|x\|=1} |\int_a^b k(u,v)\,x(v)\,dv| \le$$

$$\le \max_{a \le u \le b} \sup_{\|x\|=1} (b-a) \cdot \|k\| \cdot \|x\| =$$

$$= (b-a) \cdot \|k\|$$

and the separability of the space K, the space F is separable in the normed topology. Since F is at the same time a metric space in the topology considered, it is perfectly separable. Therefore the class of all spheres contains a countable subclass that is a base of the space F. Hence and from the validity of the relation

$$\{\omega : T(\omega,\cdot) \in \{L : \|L - L_0\| < \varepsilon\}\} =$$
$$= \bigcup_{k=1}^{\infty} \{\omega : T(\omega,\cdot) \in \{L : \|L - L_0\| \leq \varepsilon - 1/k\}\} =$$
$$= \bigcup_{k=1}^{\infty} \bigcap_{n=1}^{\infty} \{\omega : T(\omega, x_n) \in \{x : \|x - L_0(x_n)\| \leq \varepsilon - 1/k\}\}$$

for every $\varepsilon > 0$ and $L_0 \in F$, where $x_1, x_2, \dots$ is a countable dense subset of the space X, we get the statement of our theorem.

Similarly as in the deterministic case we associate to every random integral transformation T a subset of the Cartesian product $\Omega \times [-\infty, \infty]$ denoted by $\rho(T)$ and called the resolvent set of the random transformation T. To the measurability of λ-sections of the resolvent set $\rho(T)$ is devoted

T h e o r e m 17. Let all the assumptions of theorem 15 be fulfilled.

Then for every real number λ the set of those ω's for which the linear random transformation $(T - \lambda I)$ is invertible belongs to the σ-algebra $\mathfrak{S}$, i.e. for every real number

$$\{\omega : (\omega, \lambda) \in \rho(T)\} \in \mathfrak{S}.$$

P r o o f. According to the well known theorem on resolvent sets in Banach algebras the set

$$B_\lambda = \{A : A \in M, \lambda \in \rho(A)\}$$

is for every fixed λ open and therefore from the σ-algebra $\mathfrak{M}$. Hence, theorem 17 follows from theorem 16 and the relation

$$\{\omega : (\omega, \lambda) \in \rho(T)\} = \{\omega : T(\omega, \cdot) \in B_\lambda\}$$

which holds for every fixed real number λ.

A sufficient condition for the invertibility of the random transformation $(T - \lambda I)$ is given in

T h e o r e m 18. Let all the assumptions of theorem 15 be fulfilled. Let in addition the real number $\lambda \neq 0$ satisfy

$$\mu\{\omega : (b-a)\,\|k(\omega,\cdot,\cdot)\| < |\lambda|\} = 1.$$

Then the linear random transformation $(T-\lambda I)$ is invertible, i.e.

$$\mu\{\omega : (\omega,\lambda) \in \rho(T)\} = 1. \qquad /15/$$

P r o o f. From the assumptions we have for every $\omega \in \Omega$ and every $x \in X$ the inequality

$$\|T(\omega,x)\| = \|\int_a^b k(\omega,\cdot,v)\,x(v)\,dv\| \leq$$

$$\leq \int_a^b \|k(\omega,\cdot,\cdot)\|.\|x\|\,dv = (b-a).\|k(\omega,\cdot,\cdot)\|.\|x\|$$

which shows that

$$\|T(\omega,\cdot)\|_{\varepsilon} \leq (b-a).\|k(\omega,\cdot,\cdot)\|.$$

The desired conclusion then follows immediately from theorem 7.

The possibility of reaching the solution of the regression integral equation using only one realization of the stationary and ergodic kernel is furnished by

T h e o r e m 19. Let k be a mapping of the Cartesian product $\Omega \times [0,\infty)$ into the space K such that

$$\{\omega : k(\omega,t,u,v) < r\} \in \mathfrak{S}$$

for every $t \geq 0$, $(u,v) \in \Delta$, and every real number r;

$$\mu\{\omega : k(\omega,t,\cdot,\cdot) \text{ is continuous in } t\} = 1;$$

k is stationary and ergodic, i.e. in particular

$$\int_\Omega k(\omega,t,\cdot,\cdot)\,d\mu(\omega) = \int_\Omega k(\omega,0,\cdot,\cdot)\,d\mu(\omega) = q(\cdot,\cdot)$$

for every $t \geq 0$, and

$$\mu\{\omega : \lim_{t\to\infty} \|\tfrac{1}{t}\int_0^t k(\omega,s,\cdot,\cdot)\,ds - q(\cdot,\cdot)\| = 0\} = 1.$$

Let z be a stationary and ergodic transformation of the Cartesian product $\Omega \times [0,\infty)$ into the space X, i.e. in particular

$$\int_\Omega z(\omega,t)\,d\mu(\omega) = \int_\Omega z(\omega,0)\,d\mu(\omega) = z_0$$

for every $t \geq 0$, and

$$\mu\{\omega : \lim_{t\to\infty} \| \frac{1}{t}\int_0^t z(\omega,s)\,ds - z_0 \| = 0\} = 1.$$

Further, let $\lambda \neq 0$ be any real number satisfying the condition

$$\mu\{\omega : (b-a)\|k(\omega,0,\cdot,\cdot)\| < |\lambda|\} = 1 \qquad /16/$$

and denote by $x_t(\cdot)$ the unique random solution of the random operator equation

$$\int_a^b k(\cdot,0,\cdot,v)[\xi_0(\cdot)](v)\,dv - \lambda\xi_0(\cdot) = z(\cdot,0)$$

and

$$\frac{1}{t}\int_0^t\int_a^b k(\cdot,s,\cdot,v)[\xi_s(\cdot)](v)\,dv\,ds - \lambda\xi_t(\cdot) = \frac{1}{t}\int_0^t z(\cdot,s)\,ds$$

for every $t > 0$.

Then there exists a unique solution $\hat{x} \in X$ of the operator equation

$$\int_a^b q(\cdot,v)\,\xi(v)\,dv - \lambda\xi = z_0 \qquad /17/$$

and we have

$$\mu\{\omega : \lim_{t\to\infty} \|x_t(\omega) - \hat{x}\| = 0\} = 1. \qquad /18/$$

P r o o f. The existence of a unique solution follows by applying theorem 18 to the kernel q ; the relation /18/ is a consequence of theorem 11.

The next theorem forms an analog of theorem 18 for Volterra kernels.

T h e o r e m 20. Let all the assumptions of theorem 15 be fulfilled. Let in addition the kernel k satisfy the condition

$$\mu\{\omega : k(\omega,u,v) = 0\} = 1$$

for every $a \leq u < v \leq b$.

Then for every real number $\lambda \neq 0$ the linear random transformation $(T-\lambda I)$ is invertible, i.e. relation /15/ holds.

P r o o f. From our assumptions we get after several lines of computation that for every $\omega \in \Omega$ and every $n = 1, 2, \ldots$

$$\|T^n(\omega,\cdot)\| \leq |\lambda|^n \cdot \frac{(b-a)^n}{n!} \cdot \|k(\omega,\cdot,\cdot)\|^n$$

what enables us to make use of theorem 7 which yields the desired result.

Finally, we shall prove a theorem of experience type for random integral equations of Volterra type.

T h e o r e m 21. Let all the assumptions of theorem 19, with the exception of relation /16/, be fulfilled and let, in addition, the kernel k satisfy the condition

$$\mu\{\omega : k(\omega,t,u,v) = 0\} = 1$$

for every $t \geq 0$ and $a \leq u < v \leq b$.

Then there exists a unique solution $\hat{x} \in X$ of the operator equation /17/ which satisfies relation /18/.

P r o o f. The existence of a unique solution follows from theorem 20 applied to the kernel q ; relation /18/ is furnished by theorem 12.

Some tentative results

By Tore Dalenius (Stockholm), Jaroslav Hájek (Praha)
and Stefan Zubrzycki (Wroclaw)

1. Summary

We study the following problem. An isotropic plane stochastic process is observed at points making up a regular pattern. We are interested in finding patterns yielding the least limiting variance of the observed values, when the points are situated within a circle with infinitely increasing radius.

In section 4, a solution is presented for the correlation function (3.5) and some intervals of point densities. The solution is obtained by solving the related geometrical problem of covering a plane by circles in such a way that the circles mutually intersect as little as possible (see section 3). From the results obtained it follows that--in contrary to the linear case--no unique pattern of points is optimum for all convex correlation functions simultaneously. The efficiency of usually used patterns is, however, quite good.

In section 5, finally, we study a subclass of convex correlation functions of an isotropic plane process, consisting of functions which admit a spectral representation in terms of the simple correlation function (5.2).

2. Introduction

In this section, an expository survey of the background of the problem will be presented.

2.1. Applications of plane sampling. Many applications of the sampling method may be broadly described as "plane sampling." We will give some examples.

Forest surveys. In the simplest case, one might want to estimate the area of a certain country or geographical district covered by forest. Similarly, one might want to estimate the proportion of a forest area covered by a certain variety of tree. In another case, one might want to estimate the number of trees or the volume of timber in a forest area. These kinds of applilications are discussed in, that is, Matern [12].

Agricultural surveys in general. Illustrative examples of plane sampling are furnished by large-scale surveys carried out in order to estimate the cultivated area of a certain country or geographical district or the total production of a particular crop. We refer here to, for example, Mahalanobis [11].

Analogous examples are furnished by such small-scale surveys as "field experiments" restricted to a single field on a farm. There is no need for specific references on this point.

Soil surveys. One specific example concerns the estimation of the proportion of the area of ground of a certain district, which is too salty to be cultivated in the usual way; for a discussion, we refer to Sulanke [18].

Another specific example concerns the estimation of the density of worms in soil; see Finney [5].

Geological surveys proper. The method of plane

sampling has been used in order to estimate the extension, volume and other parameters characterizing such geological deposits as black and brown coal, zink deposits, etc. We refer to Zubrzycki [23].

In this connection, we want to briefly mention applications of plane sampling in connection with the construction of water power stations. In such situations, there is often need for estimating the total volume of earth to be removed (for example, as a basis for cost estimates), the total volume gravel available for construction purposes, etc.

Some other examples. The examples given relate to sampling a geographical area of some sort. The method of plane sampling is, however, applicable to sampling other kinds of areas; some references will be given.

Drápal, et al, [2] discusses the problem of estimating the proportions of the surface of cast iron composed of crystals of carbon and of iron, respectively. This problem bears a considerable resemblance to the "corpuscle problem," discussed by Wicksell [21].

Husu [9] considers the problem of estimating a parameter which characterizes the "smoothness" of the surface of a metal plate.

Faure [3] and Savelli [16] consider the problem of measuring the transparency of photographic film.

2.2. The sampling theory. The theoretical task of constructing a (probabilistic) sampling theory to cope with the problems of estimation actualized by such applications as those just discussed has long been the subject of considerable research.

By and large, the theory of "field experimentation" is the origin of the theory of plane sampling. However, from a rather early date, somewhat different paths of advancement have been taken.

In field experiments, it is often feasible to apply randomization of the experimental units; the use of randomization may be considered as a device for getting around the need for a (realistic) model of the role played by "topographic variation." In such applications of plane sampling as those discussed above, it is often desirable, for practical reasons, to use systematic sampling procedures. As a consequence, it is necessary to account for the role played by "topographic variation" by means of a (realistic) model of this variation.

The theory of linear sampling. To a large extent, the theory of plane sampling is a formal extension of the theory of "linear sampling," this term referring to sampling a one-dimensional stochastic process. Therefore, we will include some references to the linear case.

Early contributions include such papers as Osborne [14], Madow [10] and Cochran [1]. In the last-mentioned paper, it is proved that systematic sampling is, on the average, more precise than stratified sampling, provided that the correlogram is concave upwards.

Among the recent contributions we mention Hájek [6] and [7].

The theory of plane sampling. A classical contribution to the discussion of topographic variation is Smith [17]. In the field of plane sampling proper, Mahlanobis [11] may be considered the pioneer work.

Matérn [12] represents a most important contribution. In this work, Mátern demonstrates how the theory of stationary stochastic processes as developed by, that is, Khintchine and Cramér may be used in the construction of a stochastic model of topographic variation.

Quenouille [15] is another important contribution from the 1940's. Among recent contributions, we mention Whittle [19] and [20]; Williams [22]; and Zubrzycki [23] and [24].

2.3. The object of the paper. The main object of this paper is to analyze the feasibility of estimating the mean value of a certain stationary stochastic process in the plane by means of observations at points distributed over the plane according to some regular pattern.

In section 3, we develop a specific formulation of our problem. This formulation is the basis of the discussion in section 4. In section 5, finally, we present some related results.

3. Formulation of the problem of the paper

Our problem is to find a regular pattern of points which yield the least limiting variance of values associated with an isotropic plane stochastic process observed at these points. In this section, we will transform this problem into an equivalent geometrical problem, the solution of which is discussed in section 4.

To begin with, we briefly review some previously established results concerning the linear case. The typical problem can be found as follows.

To the points t of a real line there are assigned random variables $\eta(t)$, subject to the following assumptions (a) - (c).

(a) All random variables $\eta(t)$ have common expected value μ and common variance σ^2.

(b) The correlation coefficient between any two random variables $\eta(t')$ and $\eta(t'')$ depends only upon the absolute difference $|t' - t''|$, or in symbols[1]

$$R(\eta(t'), \eta(t'')) = \rho(|t' - t''|). \tag{3.1}$$

(c) The function $\rho(t)$, $t \geqq 0$, called the correlation function of the process, is continuous with $\rho(0) = 1$.

The assumptions (a) - (c) characterize the family of random variables $\eta(t)$ as a continuous stochastic process stationary to the second degree.

Suppose now that we want to estimate the mean $T^{-1}\int_0^T \eta(t)dt$ of the process by the average

$$\bar{\eta} = \frac{1}{n}(\eta(t_1) + \dots + \eta(t_n)) \tag{3.2}$$

of its values observed at n points $t_1, \dots, t_n$ selected in a given segment $0 \leqq t \leqq T$. Consider all probability methods of sampling n points such that the expected number of points selected

[1] For convenience, we number the formulas, figures, etc., serially within each section.

from any subsegment (α,β), $0 \leqq \alpha < \beta \leqq T$, is proportionate to $\beta - \alpha$. We ask how to choose these n points in order to minimize the quadratic error of estimation, defined as the expected value of $(\bar{\eta} - T^{-1}\int_0^T \eta(t)dt)^2$. This value extends over both the sampling experiment and the nature's experiment producing the process $\eta(t)$. By using an argument of Hájek's [7] involving a kind of spectral representation of the correlation function of the process with respect to a one parameter family of properly chosen simple correlation functions--it follows that if the correlation function is convex, then the best method of choosing the n points $t_1 \ldots t_n$ is to select them equidistantly.

We now try to generalize the discussion of the linear case to the case of a plane. Thus, we consider a family of random variables $\eta(p)$ assigned to the points p of an Euclidean plane, for which the following generalizations (a') - (c') of the previously given assumptions (a) - (c) are fulfilled:

(a') all random variables $\eta(p)$ have common expected value μ and common variance σ^2;

(b') the coefficient of correlation between any two random variables $\eta(p)$ and $\eta(q)$ depends only on the vector joining the points p and q, or in symbols

$$R\{\eta(p), \eta(q)\} = \rho(q - p) \tag{3.3}$$

where q - p is the vector difference between q and p.

(c') The correlation function $\rho(p)$ is a continuous function of p with $\rho(0) = 1$, where 0 is the zero-vecto

In the sequel we shall be concerned with processes which are, in addition, isotropic. That is to say, the correlation function depends only on the length $u = |p - q|$ of of the vector p - q

$$\rho(p - q) = \rho(|p - q|) = \rho(u) \tag{3.4}$$

that is, the correlation function is a function of one real variable. In what follows, we refer to this function $\rho(u)$ as the correlation function.

In the linear case, it turned out that the best method of sampling n points is to select them equidistantly. It is, however, difficult to generalize this result in a straightforward manner to the case of a plane; it is not obvious which domains in the plane can replace the segment $0 \leq t \leq T$. Therefore, we have looked for such regular allocations that can be dealt with by means of limiting theorems, relating to increasing domains. This approach eliminates the troublesome boundary effect from the problem.[1] On the other side, it introduces some problems of a purely geometrical nature which we are now going to discuss.

As shown in Zubrzycki [23], we can construct a two-dimensional continuous and isotropic stationary stochastic process $\eta(p)$, the correlation function of which is $\rho(u) = r(u/a)$

$$r(\tfrac{u}{a}) = \frac{2}{\pi} \{\arc \cos \frac{u}{a} - \frac{u}{a} \sqrt{1 - (\frac{u}{a})^2}\} \quad \text{for } 0 \leq u \leq 1, \tag{3.5}$$

$$= 0 \text{ otherwise}$$

and a is a positive constant.

1 An alternative device to cope with the boundary effect would be to define the stochastic process $\eta(p)$ on the surface of a sphere or a torus.

In this analytic form, the correlation function does not reveal its most important feature: the value of $r(u/a)$ for a given u can be computed as

$$(3.6) \qquad r\left(\frac{u}{a}\right) = \frac{|K_1 \cap K_2|}{\pi \frac{a^2}{4}}$$

that is, $r(u/a)$ equals the ratio of the area $|K_1 \cap K_2|$ of the common part of two circles K_1 and K_2 with radius $\alpha/2$ and centers p_1 and p_2 at distance $u = |p_1 - p_2|$ to the area $\pi a^2/4$ of such a circle. The geometrical interpretation is illustrated in Fig. 3.1.

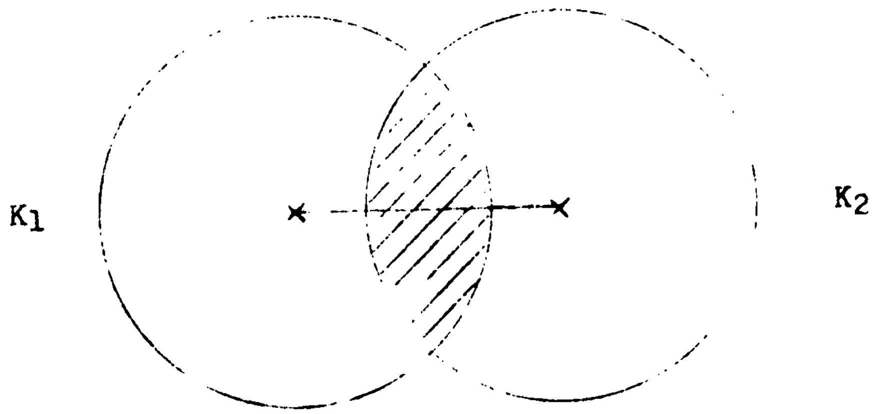

Fig. 3.1. Geometrical interpretation of the value of the correlation function given by (3.5).

Suppose now that we have selected n points $p_1 \ldots p_n$ in a plane and want to estimate the mean value μ of the process[1] by the average

---

1 The mean value μ may be interpreted as $|D|^{-1} \int_D \eta(p)dp$ for a domain D with infinitely large area $|D|$.

$$\bar{\eta} = \frac{1}{n} \{\eta(p_1) + \ldots + \eta(p_n)\} \tag{3.7}$$

of the values $\eta(p_1) \ldots \eta(p_n)$ at these points. This average $\bar{\eta}$ is an unbiased estimate of μ. Obviously,

$$\operatorname{Var} \bar{\eta} = \frac{\sigma^2}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \rho(|p_i - p_j|) . \tag{3.8}$$

The double sum of the right member of (3.8) is, by virtue of the geometrical interpretation of (3.5) equal to the ratio of the sum of the areas of the common parts of all n^2 possible pairs of circles K_i with radius $a/2$ and centers p_i, $i = 1 \ldots n$, to the area of such a circle, or in symbols

$$\operatorname{Var} \bar{\eta} = \frac{\sigma^2}{n^2} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} |K_i \cap K_j|}{\pi a^2/4} . \tag{3.9}$$

We now ask which distribution of points in the plane with a given plane density yields (in the limit) the least variance. What we said in connection with the correlation function given by (3.5) shows that this problem is uniquely related to certain questions concerning distributions of circles. Let us consider a sequence of points $p_1 \ldots p_n$ in the plane and define the density d, if it exists, as

$$d = \lim_{R \to \infty} \frac{1}{\pi R^2} \operatorname{Card} \{i: p_i \in K(0, R)\} , \tag{3.10}$$

that is, as a limit of a ratio of the number of points p_i in a circle $K(0, R)$ of radius R and center 0 to the area πR^2 of this circle, when R tends to infinity. Of course, this limit does not depend on the choice of center O. Then, given a sequence $p_1, p_2, \ldots$ with density d, let us call a <u>limiting variance</u> a limit

$$(3.11) \qquad \sigma_\infty^2 = \lim_{R \to \infty} n_R \cdot \operatorname{Var} \bar{\eta}_R$$

where $n_R = \operatorname{Card}\{i: p_i \in K(0, R)\}$ and $\bar{\eta}_R$ is the average of those variables $\eta(p_i)$ for which p_i is in $K(0, R)$.

Let us introduce two more definitions. Given in a plane a sequence of circles $K_1, K_2, \ldots$ with radius $a/2$, the centers $p_1, p_2, \ldots$ of which have a given density d, we will call a <u>mean covering</u>, in short C', a limit

$$(3.12) \qquad C' = \lim_{R \to \infty} \frac{1}{\pi R^2} \sum_i |K_i \cap K(0, R)|$$

and a <u>mean double covering</u>, in short C", a limit

$$(3.13) \qquad C'' = \lim_{R \to \infty} \frac{1}{\pi R^2} \sum_i \sum_j |K_i \cap K_j \cap K(0, R)| \quad .$$

In this sum, the case $i = j$ is not excluded. Of course, the mean covering C' of the circles $K_1, K_2, \ldots$ and density d of their centers $p_1, p_2, \ldots$ are related by the equality $C' = |K_1| d$. As a consequence, we may use C' as our measure of density of centers in comparisons where $|K_1|$ is kept constant.

Now it is clear that for stochastic processes $\eta(p)$ with correlation functions given by (3.5) the search for a sequence of points with a prescribed density which yields the minimum limiting variance is equivalent with the search for a corresponding sequence of circles yielding the minimum mean double covering. Moreover, if there exists such a sequence of points which would realize the minimum limiting variance for all positive values of a in the correlation function given by (3.5), then it would be the best sequence also for a process with a correlation function given by

$$\rho(u) = \int_0^\infty r(u/a)\, dF(a) \tag{3.14}$$

where $F(a)$ is a distribution function with $F(0) = 0$.

It is of considerable interest to find out which correlation functions admit the representation given by (3.14). Unfortunately, there do not exist sequences of points which yield minimum mean double covering simultaneously for all values of a in (3.5), as will be shown later on in this paper.

4. Optimal nets of points

Consider a sequence of congruent circles $K_1, K_2 \ldots$ with centers $p_1, p_2 \ldots$ respectively. Let us denote the indicator function of K_i by $k_i(p)$, that is, let us put

$$k_i(p) = \begin{cases} 1 & \text{for } p \in K_i, \\ 0 & \text{otherwise} \end{cases} . \tag{4.1}$$

Moreover, we put

$$k(p) = \sum_i k_i(p) \quad . \tag{4.2}$$

In other words, $k(p)$ is equal to the number of circles covering p. In terms of these functions, the definitions of the mean covering C' and the mean double covering C" given in section 3 take on the forms

$$(4.3) \qquad C' = \lim_{R\to\infty} \frac{1}{\pi R^2} \sum_i \int_{K(0,R)} k_i(p)dp =$$

$$= \lim_{R\to\infty} \frac{1}{\pi R^2} \int_{K(0,R)} \sum_i k_i(p)dp =$$

$$= \lim_{R\to\infty} \frac{1}{\pi R^2} \int_{K(0,R)} k(p)dp;$$

$$(4.4) \qquad C'' = \lim_{R\to\infty} \frac{1}{\pi R^2} \sum_i \sum_j \int_{K(0,R)} k_i(p)k_j(p)dp =$$

$$= \lim_{R\to\infty} \frac{1}{\pi R^2} \int_{K(0,R)} \sum_i \sum_j k_i(p)\, k_j(p)dp =$$

$$= \lim_{R\to\infty} \frac{1}{\pi R^2} \int_{K(0,R)} k^2(p)dp.$$

Our problem is to determine such sequences of congruent circles with a fixed mean covering for which the mean double covering attains its minimum. We are now going to prove an inequality from which it follows that a sufficient condition for a sequence of circles to have this minimum property is that the set of values of the function k(p) consists of two consecutive integers. It is the content of the following.

<u>Lemma</u> 4a. <u>For any sequence of circles with mean covering</u> C', <u>the following inequality</u>

$$(4.5) \qquad C'' \geqq \{2[C'] + 1\}\, C' - [C']\, \{[C'] + 1\}$$

holds, where $[C']$ is the integral part of C'; equality holds if $k(p)$ takes only the values $[C']$ and $[C'] + 1$.

Proof. Clearly, we have

$$\frac{1}{\pi R^2}\int_{K(0,R)} k^2(p)dp = \left\{\frac{1}{\pi R^2}\int_{K(0,R)} k(p)dp\right\}^2 \tag{4.6}$$

$$- \left\{\frac{1}{\pi R^2}\int_{K(0,R)} k(p)dp - [C'] - \frac{1}{2}\right\}^2$$

$$+ \frac{1}{\pi R^2}\int_{K(0,R)} \left(k(p) - [C'] - \frac{1}{2}\right)^2 dp \quad .$$

As always,

$$\left(k(p) - [C'] - \frac{1}{2}\right)^2 \geqq \frac{1}{4} \quad , \tag{4.7}$$

we conclude that

$$\frac{1}{\pi R^2}\int_{K(0,R)} k^2(p)dp \geqq \left\{\frac{1}{\pi R^2}\int_{K(0,R)} k(p)dp\right\}^2 \tag{4.8}$$

$$- \left\{\frac{1}{\pi R^2}\int_{K(0,R)} k(p)\, dp - [C'] - \frac{1}{2}\right\}^2 + \frac{1}{4} \, .$$

For $R \longrightarrow \infty$, we get

$$C'' \geqq C'^2 - \left\{ C' - [C'] - \frac{1}{2}\right\}^2 + \frac{1}{4} \tag{4.9}$$

and this is an alternative form of (4.5). Now if $[C']$ and $[C'] + 1$ are the only values of $k(p)$, then (4.7) and consequently (4.5) become equalities. This proves the lemma.

We will now describe some sequences of circles minimizing the mean double covering. We confine ourselves to the case where the centers of the circles form nets composed of congruent figures such as triangles, squares, etc. This will enable us to compute the mean covering and the mean double covering from a single mesh of a net and we shall exploit this possibility. The minimum property will follow by our lemma, as the function k(p) will take only two consecutive integers as its values in our examples. We arrange these examples by increasing values of C'

Case I. If

$$C' \leqq \frac{\pi}{2\sqrt{3}} \doteq 0.907 \ldots$$

then the net of equilateral triangles[1] has the optimal property. This situation is illustrated in Fig. 4.1 and Fig. 4.2.

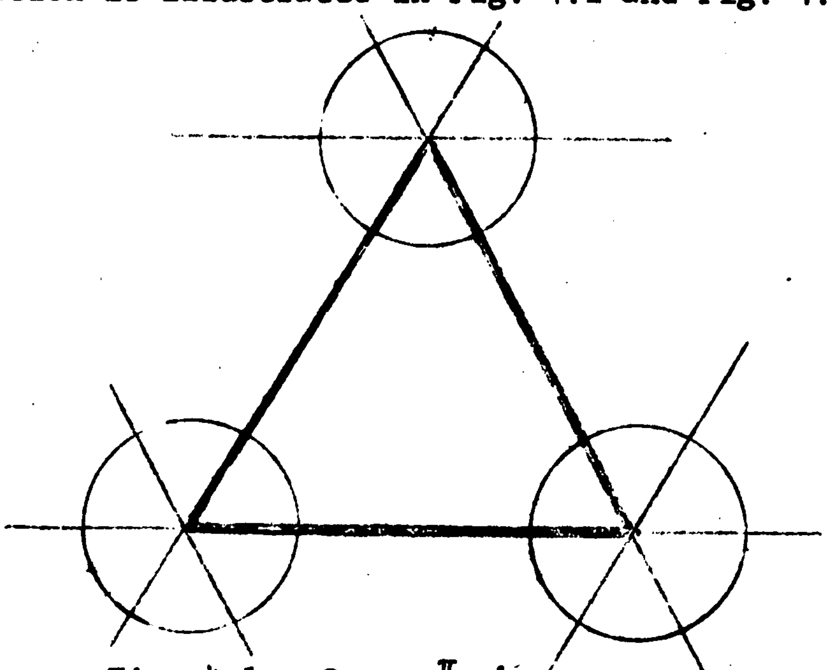

Fig. 4.1. $C' < \frac{\pi}{2\sqrt{3}} \doteq 0.907\ldots$.

---

1 The pattern may just as well be referred to as a "net of rhombuses."

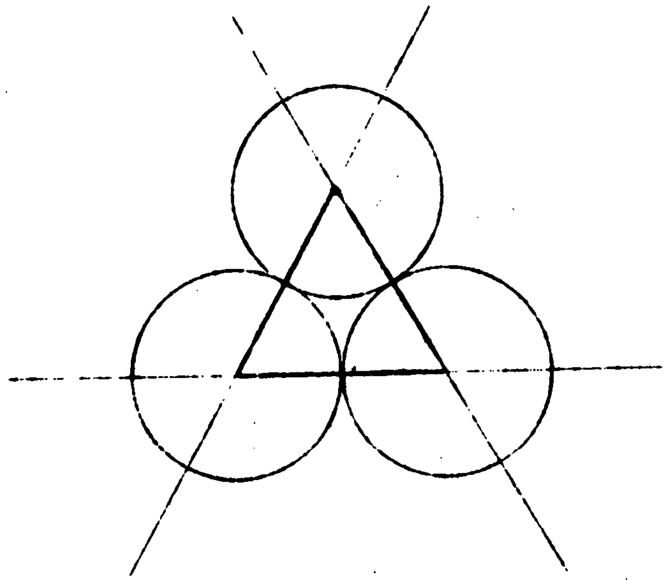

Fig. 4.2. $C' = \frac{\pi}{2\sqrt{3}} \doteq 0.907 \ldots$.

For the purpose of illustration, we will present here the details of the computations of C' and C''.

In Fig. 4.1, we put the radius R of the circles equal to 1, and the side of the triangle equal to $s \geq 2$. Thus, the area of the triangle is

$$A = \frac{s^2 \sqrt{3}}{4} \quad .$$

The circles divide this area into four parts

$$A = T_0 + 3T_1 = A_o + A_1 \ .$$

The meaning of T_o and T_1 is shown in Fig. 4.3.

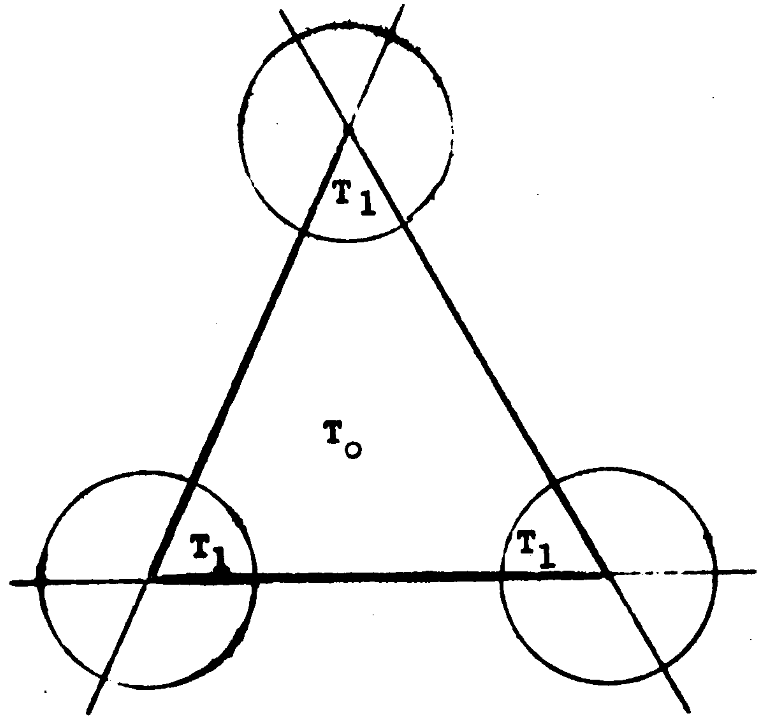

Fig. 4.3.

In T_o, $k(p) = 0$, while in T_1, $k(p) = 1$. Now

$$C' = \frac{1}{A} \{ \int_{A_o} 0 \, dp + \int_{A_1} dp \} = \frac{1}{A} \int_{A_1} 1 \, dp.$$

Thus,

$$C' = \frac{4}{s^2 \sqrt{3}} \{3 \cdot \frac{\pi}{6}\} = \frac{2\pi}{s^2\sqrt{3}} .$$

For $s > 2$, $C' < \pi/2\sqrt{3} \doteq 0.907 \ldots$ while for $s = 2$, $C' = \pi/2\sqrt{3} \doteq 0.907 \ldots$. Moreover, $C'' = C'$ as

$$C'' = \frac{1}{A} \{ \int_{A_o} 0^2 \, dp + \int_{A_1} 1^2 \, dp \} = C'.$$

We will now compare this value of C" with the corresponding value of C" for a net of squares having the same value of C', and therefore, representing the same point density.

We put the side of the square equal to x. Drawing the four circles with radius $R = 1$, we get, say

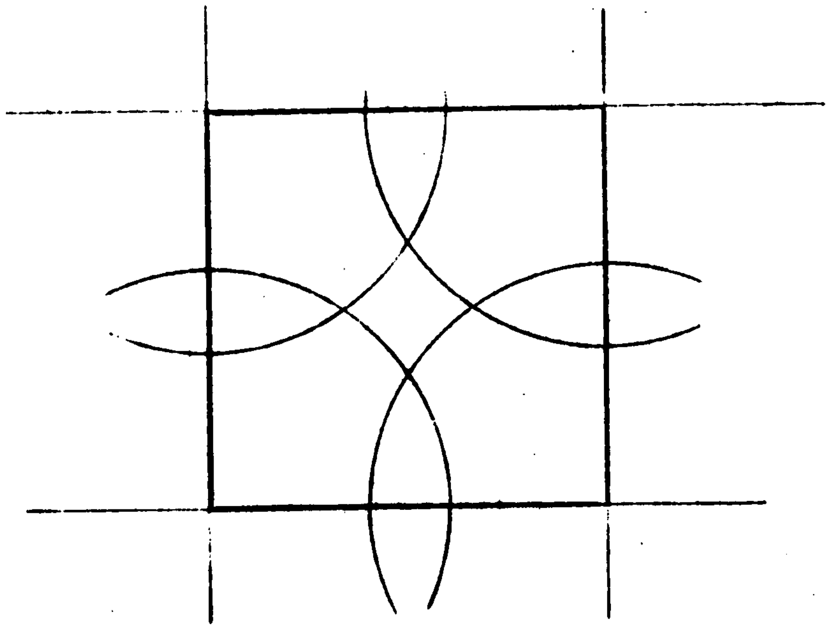

Fig. 4.4.

Now

$$C' = \frac{1}{A}\left\{\int_{A_o} 0\, dp + \int_{A_1} 1\, dp + \int_{A_2} 2\, dp\right\} = \frac{\pi}{x^2} .$$

From $C' = \pi/2\sqrt{3}$, we get

$$x^2 = \sqrt{12};\ x = \sqrt[4]{12} \doteq 1.86 < 2R.$$

If A_2 stands for the area of that portion of the square, where $k(p) = 2$, we get

$$A_2 = 8\left\{ \frac{v}{360} \cdot \pi - \frac{1}{2} \cdot y \cdot \frac{\sqrt[4]{12}}{2} \right\} = 8\left\{\frac{v}{360}\, \pi - \frac{1}{4} \sin v \sqrt[4]{12}\right.$$

v and y having the meaning indicated in Fig. 4.5.

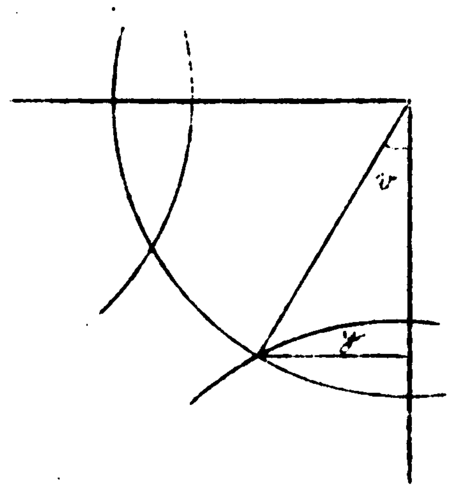

Fig. 4.5.

Carrying out the computations gives $A_2 = 0.14$. Clearly $A_1 = \pi - 2A_2 = 2.86$. Thus

$$C' = \frac{1}{\sqrt{12}} [2.86 + 2 \cdot 0.14] = 0.907\ldots$$

(as it should be), and

$$C'' = \frac{1}{\sqrt{12}} \ [2.86 + 2^2 \cdot 0.14] = 0.988 \ldots > 0.907\ldots$$

Case II. If

$$0.907\ldots \doteq \frac{\pi}{2\sqrt{3}} \leqq C' \leqq \frac{2\pi}{3\sqrt{3}} \doteq 1.209 \ldots$$

then the net of equilateral triangles is still optimal. See Fig. 4.6.

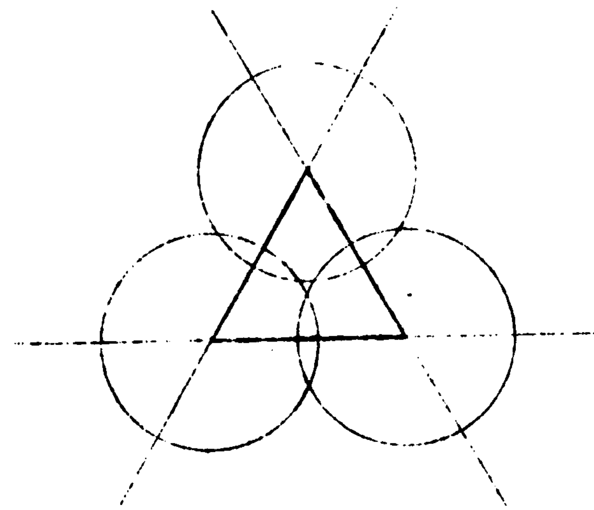

Fig. 4.6. $0.907 \ldots \doteq \frac{\pi}{2\sqrt{3}} < C' < \frac{2\pi}{3\sqrt{3}} \doteq 1.209.$

However, in this case k(p) has three values: 0, 1, and 2, so that our lemma 4a does not apply. The optimality of the net in question is a consequence of a known inequality (Toth Fejes [4], Chapter III, paragraph 8, inequality (3), p. 80) from which it follows that among all convex hexagons of a given area and circles of a given area, the maximum possible area of a common part of a hexagon and circle is reached, if the hexagon is equilateral and the circle is concentric with it.

The statement of case II follows, if we apply the quoted inequality to cells which are formed by attaching each point of a plane to the nearest circle center. The above-mentioned inequality is of its greatest interest when the mean covering is in the range indicated in Case II. Let us note, however, that it implies also the statement in Case I to the effect that the circles should be disjoint.

Case III. If

$$1.209 \ldots \doteq \frac{2\pi}{3\sqrt{3}} \leq C' \leq \frac{2\pi}{2+\sqrt{3}} \doteq 1.684\ldots$$

then the net of isosceles triangles is optimal.

This is seen as follows. We start with the situation drawn in Fig. 4.7.

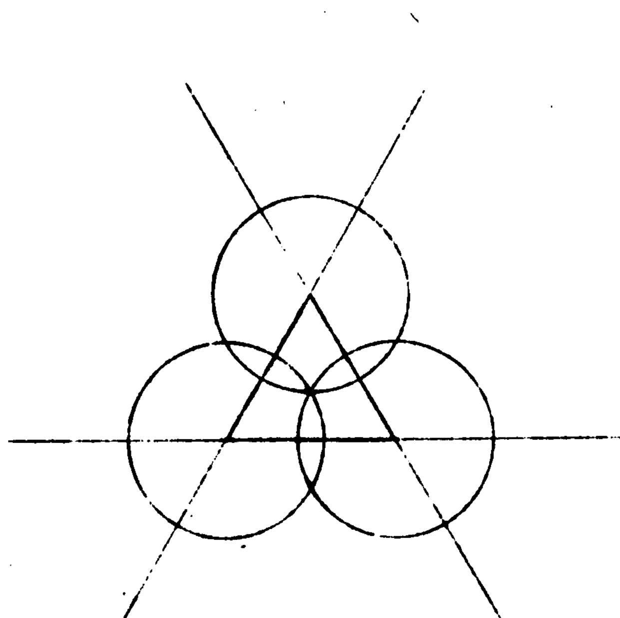

Fig. 4.7. $C' = \frac{2\pi}{3\sqrt{3}} \doteq 1.209 \ldots$.

We then increase the mean covering without spoiling the property of k(p) of having only two consecutive integers as values, letting the base of the triangle diminish and the height of it enlarge, so that the three circles intersect still in one point. We can continue this procedure until the length of the base becomes equal to the radius of our circles, as shown in Fig. 4.8.

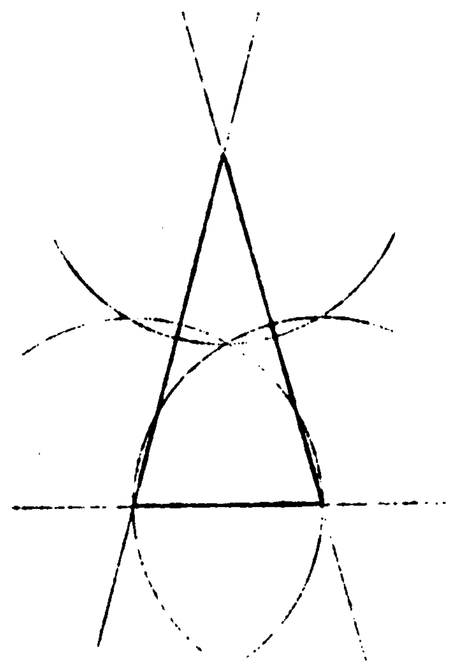

Fig. 4.8. $C' = \frac{2\pi}{2+\sqrt{3}} \doteq 1.684 \ldots$.

Case IV. If

$$1.571\ldots \doteq \frac{\pi}{2} \leqq C' \leqq \frac{\pi}{\sqrt{3}} \doteq 1.814\ldots$$

a net of rectangles has the optimal property.

We start with a net of squares and circles intersecting in the centers of the squares as indicated in Fig. 4.9.

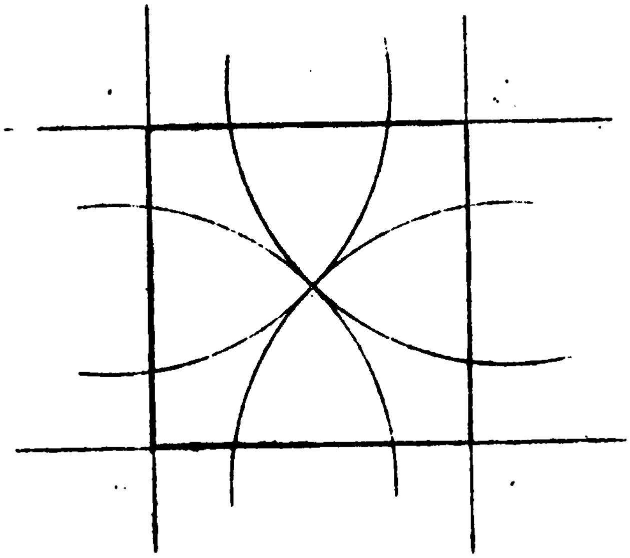

Fig. 4.9. $C' = \frac{\pi}{2} \doteq 1.571\ldots$.

This corresponds to $C' = \pi/2 \doteq 1.571 \ldots$. We then enlarge the mean covering by lengthening two sides of the square and shortening two others, while the circles still intersect in the middle, as shown in Fig. 4.10, where the extreme situation is drawn. The radius of the circles is then equal to the shorter side of the rectangle. This corresponds to $C' = \pi/\sqrt{3} \doteq 1.814\ldots$.

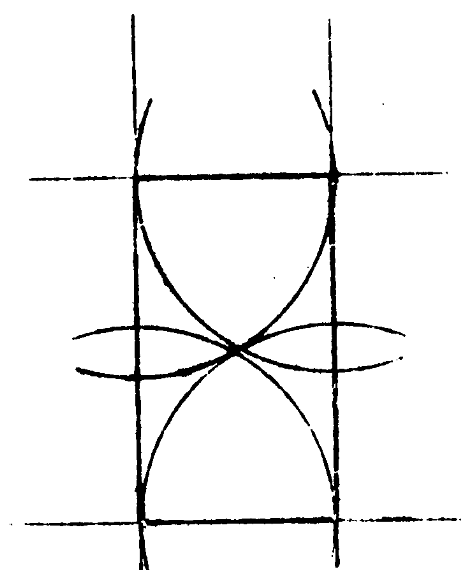

Fig. 4.10. $C' \frac{\pi}{\sqrt{3}} \doteq 1.814\ldots$.

We note that the intervals for C' corresponding to case III and case IV respectively overlap. The nature of this situation will be somewhat elucidated here.

Instead of considering the pattern with which we start in case III as made up by a net of equilateral triangles, we think of this pattern in terms of rhombuses, with the base angle $v = 60^{\circ}$, corresponding to $C' = 2\pi/3\sqrt{3} \doteq 1.209 \ldots$. If we increase v to $v = 90^{\circ}$, that is, we change the rhombuses into squares, C' will increase to $C' = \pi/2 \doteq 1.571 \ldots$. Thereafter, by "stretching" the squares into rectangles, we may further increase C' to $C' = \pi/\sqrt{3} \doteq 1.814 \ldots$.

Case V.

$$2.418 \ldots \doteq \frac{4\pi}{3\sqrt{3}} \leqq C' \leqq \frac{16\pi}{7\sqrt{7}} \doteq 2.714 \ldots$$

the optimal property is possessed by a net of hexagons which have two perpendicular axes of symmetry and can be inscribed in a circle; in general, they are not equilateral.

We start with a net of equilateral, congruent hexagons and place the centers of circles at the vertices, the radius of the circles being equal to the side of the hexagons. In this case $C' = 4\pi/3\sqrt{3} \doteq 2.418 \ldots$. This case is shown in Fig. 4.11.

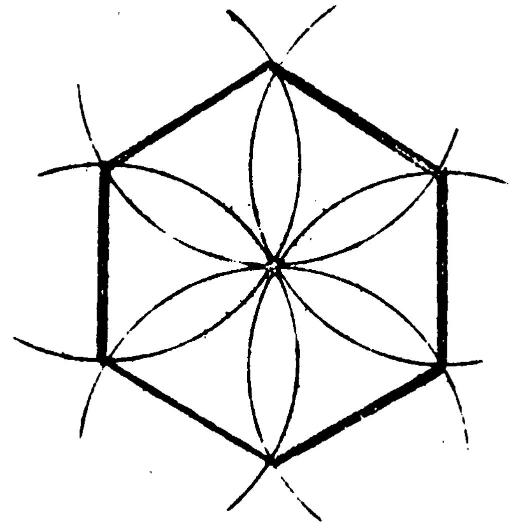

Fig. 4.11. $C' = \frac{4\pi}{3\sqrt{3}} \doteq 2.418\ldots$.

We let C' increase without spoiling the property of k(p) of having only two consecutive integers as values, by narrowing suitably our hexagons. We can continue this procedure until the circles corresponding to the vertices of neighboring hexagons come in touch. Fig. 4.12 shows the extreme situation, which corresponds to $C' = 16\pi/7\sqrt{7} \doteq 2.714\ldots$

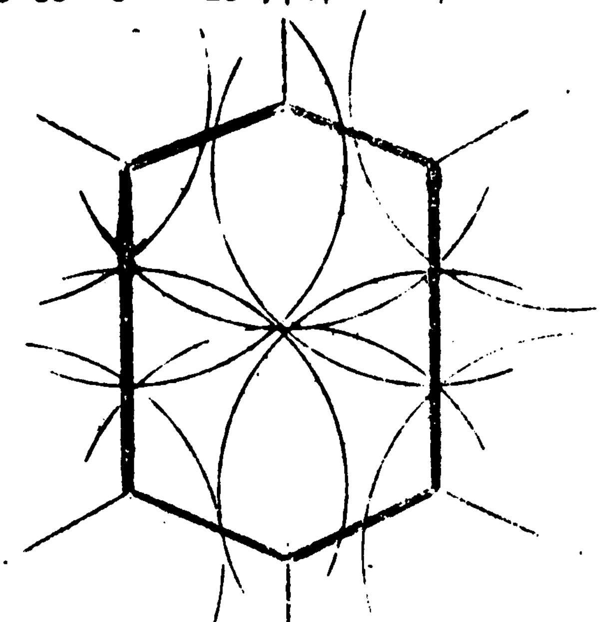

Fig. 4.12. $C' = \frac{16\pi}{7\sqrt{7}} \doteq 2.714\ \ldots$.

Case VI. If

$$C' = \frac{2\pi}{\sqrt{3}} \doteq 3.628\ldots$$

the net of equilateral triangles has the optimal property. This pattern is represented in Fig. 4.13.

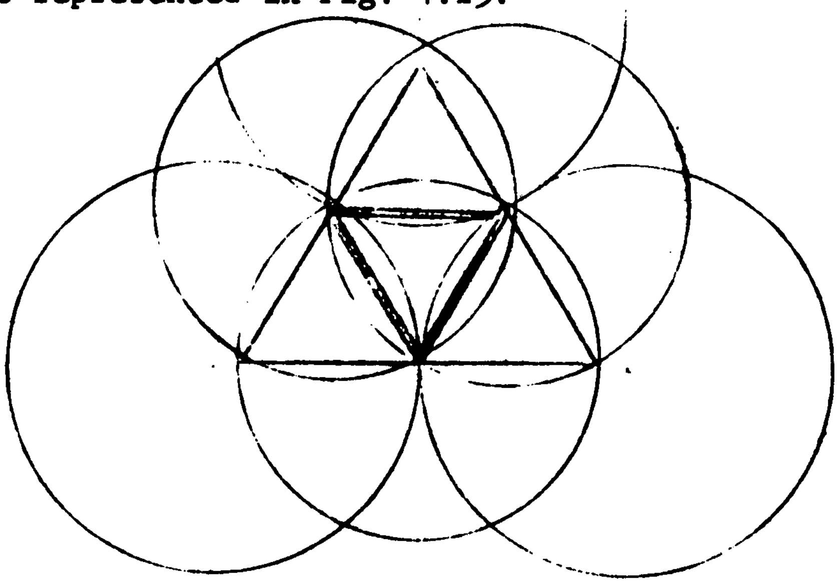

Fig. 4.13. $C' = \frac{2\pi}{\sqrt{3}} \doteq 3.628\ \ldots$.

We may summarize the previous findings as follows. By means of lemma 4a, optimal regular nets of sample points were found for the following values of C'

CASE	VALUE OF C'
I	$C' \leq \frac{\pi}{2\sqrt{3}} \doteq 0.907\ldots$
III + IV	$1.209\ \ldots \doteq \frac{2\pi}{3\sqrt{3}} \leq C' \leq \frac{\pi}{\sqrt{3}} \doteq 1.814\ldots$
V	$2.418\ \ldots \doteq \frac{4\pi}{3\sqrt{3}} \leq C' \leq \frac{16\pi}{7\sqrt{7}} \doteq 2.714\ldots$ [1]
VI	$C' = \frac{2\pi}{\sqrt{3}} \doteq 3.628\ldots$

1 On intuitive grounds, I suspect that for case V we should have

$$2.418\ \ldots \doteq \frac{4\pi}{3\sqrt{3}} \leq C' \leq \frac{3\pi}{2\sqrt{3}} \doteq 2.721\ \ldots$$

I have not had the opportunity to discuss this footnote with

We observe that lemma 4a does not provide solutions for values of C' close to 1, 2, and 3 respectively. We conjecture that the same holds true for sufficiently large values of C'. This leads us to formulate the following

Problem 4.1. Determine the range of values of C' for which

$$C'' = \{2\,[C'] + 1\}\,C' - [C']\{[C'] + 1\}\ .$$

For solving this problem, Heppes [8] should prove valuable.

It is of some interest to compare, for given values of C', the corresponding values of C" for different nets of sample points. In the following Table 4.1, we present such a comparison

TABLE 4.1

The Relation Between C" and the Pattern of the Net of Points for Some Values of C'.

C'		[C']	C" when the pattern of the net of points is		
exact value	Num. value		equilateral triangles	squares	equilateral hexagons
$\frac{\pi}{2\sqrt{3}}$	0.907	0	0.907	0.988	1.165
$\frac{2\pi}{3\sqrt{3}}$	1.209	1	1.627	1.695	1.870
$\frac{\pi}{2}$	1.571	1	2.861	2.712	2.814
$\frac{4\pi}{3\sqrt{3}}$	2.418	2	6.395	6.389	6.110
$\frac{16\pi}{7\sqrt{7}}$	2.714	2	7.711	7.936	7.736

The cases discussed above and the examples given in Table 4.1 show that the net of equilateral triangles, is not, in general, the optimal net. In other words, this net does not yield the minimum limiting variance for an isotropic stochastic process in a plane with correlation function given by (3.5). The differences are, however, rather small and their sign alternates, if C' increases. Therefore, we state the following

Problem 4.2. Is it true that the net of equilateral triangles yields the minimum limiting variance for a stationary isotropic stochastic process with exponential correlation function

The affirmative answer to this problem would imply the optimality of the net of equilateral triangles for all isotropic processes with any completely monotone correlation function. Numerical examples contained in Matérn [12] give a certain support to such an affirmative answer. Matérn compares the limiting variances for an isotropic process with correlation function e^{μ}, if triangular, square, and hexagonal nets of some chosen densities are used. The effect is that in all cases triangular net proved to be somewhat better than the other ones.

Finally, we mention without formal proof the almost obvious relation concerning the limiting behavior of the mean double covering. Let us write it down as

Lemma 4b. For any regular net of points the ratio of the mean double covering C'' and the square of the mean covering C'^2 tends to unity when the mean covering grows over all bands; in symbols

(4.10) $$\lim_{C' \to \infty} \frac{C''}{C'^2} = 1.$$

5. A class of planar isotropic correlation functions[1]

We shall discuss the class of convex planar correlation functions admitting representation

$$(5.1)\qquad \rho(u) = \int_0^\infty r\left(\frac{u}{a}\right)\, dF(a)$$

where $F(a)$ is a distribution function with $F(0+) = 0$, and

$$(5.2)\qquad r(u) = \frac{2}{\pi}\left(\arccos u - u\sqrt{1-u^2}\right) \quad \text{if } 0 \leqq u \leqq 1$$
$$= 0 \qquad \text{if } u > 1 \;.$$

Clearly,

$$(5.3)\qquad r'(u) = -\frac{4}{\pi}\sqrt{1-u^2} \qquad \text{if } 0 \leqq u < 1$$
$$= 0 \qquad \text{if } u > 1 \;,$$

and

$$(5.4)\qquad r''(u) = \frac{4}{\pi}\frac{u}{\sqrt{1-u^2}} \qquad \text{if } 0 \leqq u < 1$$
$$= 0 \qquad \text{if } u > 1.$$

Lemma 5a.

$$\int_0^\infty \frac{1}{u^2}\, r''\left(\frac{a}{u}\right) r''\left(\frac{u}{\mu}\right) du = \frac{1}{\mu}\,\frac{8}{\pi} \qquad \text{if } \mu > a > 0$$
$$= 0 \qquad \text{if } a > \mu.$$

Proof. The case $a > \mu$ is clear. If $\mu > a$, we have, in view of (5.4)

---
1 This paragraph is due to Hajek and Zubrzycki.

$$\int_0^\infty \frac{1}{u^2} \, r'' \left(\frac{a}{u}\right) r''\left(\frac{u}{\mu}\right) du =$$

$$= \left(\frac{4}{\pi}\right)^2 \int_a^\mu \frac{a \, du}{u\sqrt{(u^2 - a^2)(\mu^2 - u^2)}} =$$

$$= \left(\frac{4}{\pi}\right)^2 \left[- \frac{1}{\mu} \arccos \frac{\sqrt{1 - \frac{a^2}{u^2}}}{\sqrt{1 - \frac{a^2}{\mu^2}}} \right]_a^\mu =$$

$$= \frac{1}{\mu} \left(\frac{4}{\pi}\right)^2 \frac{\pi}{2} = \frac{1}{\mu} \frac{8}{\pi} \quad .$$

Theorem 5.1. The correlation functions $\rho(u)$ of type given by (5.1) are characterized by the following properties:

(d) $\rho(u)$ is continuous, convex, and with $\rho(\infty) = 0$.

(e) $\rho'(u)$ is absolutely continuous.

(f) $\int_0^\infty (1/u^2) \, r''(a/u) \, \rho''(u) \, du$ is a nondecreasing function of a.

The functions $F(a)$ and $\rho''(u)$ are linked by the following inversion formulas:

$$dF(a) = - \frac{\pi}{8} a^3 \, d \int_0^\infty \frac{1}{u^2} \, r''\left(\frac{a}{u}\right) \rho''(u) \, du \; . \tag{5.6}$$

$$\rho''(u) = \int_0^\infty \frac{1}{a^2} \, r''\left(\frac{u}{a}\right) dF(a) \; . \tag{5.7}$$

The condition (f) is fulfilled, for example, if $(\rho''(u)/u)$ is a nondecreasing function of u, which means, provided that $\rho''(u)$ is absolutely continuous, that

$$\rho''(u) - u\rho'''(u) \geqq 0 \quad u \geqq 0. \tag{5.8}$$

If (5.8) holds, then F(a) is absolutely continuous and therefore

$$\frac{dF(a)}{da} = -\frac{1}{2} a^2 \int_0^{\pi/2} \rho'''\left(\frac{a}{\sin\theta}\right) \frac{d\theta}{\sin^2\theta} . \tag{5.9}$$

Proof. The property (d) easily follows from the corresponding property of correlation functions $r(u/a)$. The property (e), and simultaneously the relation given by (5.7), will be proved, if we show that the indefinite integral of the right side of (5.7) equals $\rho'(u)$. Now,

$$\int_u^\infty \int_0^\infty \frac{1}{a^2} r''\left(\frac{u}{a}\right) dF(a)\, du = \int_0^\infty \frac{1}{a}\left[\int_u^\infty \frac{1}{a} r''\left(\frac{u}{a}\right) du\right] dF(a)$$

$$= -\int_0^\infty \frac{1}{a} r'\left(\frac{u}{a}\right) dF(a) = -\rho'(u$$

The change of integration order is justified, since $r''(u) \geqq 0$ (Fubini's theorem). The last identity follows from (5.1) by differentiation under the integral sign, which is justified, since $(1/a)\, r'(u/a)$ is uniformly bounded for $u > \varepsilon > 0$.

In view of (5.7), we have (by using Fubini's theorem again

$$\int_0^\infty \frac{1}{u^2} r''\left(\frac{a}{u}\right) \rho''(u)\, du = \int_0^\infty \frac{1}{u^2} r''\left(\frac{a}{u}\right) \int_0^\infty \frac{1}{\mu^2} r''\left(\frac{u}{\mu}\right) dF(\mu)\, du =$$

$$= \int_0^\infty \frac{1}{\mu^2}\left[\int_0^\infty \frac{1}{u^2} r''\left(\frac{a}{u}\right) r''\left(\frac{u}{\mu}\right) du\right] dF(\mu)$$

which gives, in accordance with (5.5)

(5.10) $$\int_0^\infty \frac{1}{u^2} r''\left(\frac{a}{u}\right) \rho''(u)\, du = \frac{8}{\pi}\int_a^\infty \frac{1}{\mu^3}\, dF(\mu) .$$

The last relation, however, is equivalent with (5.6).

Now, assume that the correlation function $\rho(u)$ fulfills the conditions (d), (e) and (f), and consider the function $F(a)$ given by (5.6). In view of the property (f), $F(a)$ will be nondecreasing.

The fact that the total variation of $F(a)$ equals 1 follows from the subsequent lemmas 5b and 5c and from the following relations:

(5.11) $$\int_0^\infty dF(a) = -\frac{\pi}{8}\int_0^\infty a^3\, d\int_0^\infty \frac{1}{u^2} r''\left(\frac{a}{u}\right) \rho''(u)\, du\, da =$$

$$= \left[-\frac{\pi}{8} a^3 \int_0^\infty \frac{1}{u^2} r''\left(\frac{a}{u}\right) \rho''(u) du\right]_0^\infty +$$

$$+ \frac{3\pi}{8}\int_0^\infty \left[\int_0^\infty \frac{a^2}{u^2} r\left(\frac{a}{u}\right) da\right] \rho''(u)\, du =$$

$$= \int_0^\infty u\rho''(u)\, du = -\int_0^\infty \rho'(u)\, du =$$

$$= \rho(0) - \rho(\infty) = 1.$$

It remains to show that $\rho'(u)$ is uniquely determined by the relation (5.6), that is, that $\rho(u)$ coincides with the correlation function obtained from (5.1). However, (5.6) is equivalent to (5.10), if rewritten in the following form:

(5.12) $$\frac{8}{\pi}\int_0^\infty \frac{1}{\mu^3}\, dF(\mu) = \int_0^\infty \frac{1}{u^2} r''\left(\frac{a}{u}\right)\, d\,\rho'(u).$$

Comparing (5.12) with (5.7), we can see that the $\rho'(u)$ may be determined from $(8/\pi) \int_a^{\infty}(1/\mu^3)dF(\mu)$ in the same way as $F(a)$ has been determined from $\rho''(u)$. (The fact that $\rho'(u)$ may not have finite variation is irrelevant.)

Before proceeding to the rest of the proof, we observe that by substituting $u = (a/\sin\theta)$ into (5.6) we get

$$(5.13) \qquad dF(a) = -\frac{1}{2}a^3 \, d\int_0^{\pi/2} \frac{\rho''(\frac{a}{\sin\theta})}{\frac{a}{\sin\theta}} \, d\theta \quad .$$

From this form it is easily seen that the condition (f) is fulfilled if $(\rho''(u))/u$ is a nondecreasing function of u, or, more especially, if (5.8) holds (notice that $[(\rho''(u))/u]' = -1/u^2 \{\rho''(u) - u\rho'''(u)\}$). Now, if (5.8) holds, we can differentiate in (5.13) under the integral sign (Fubini's theorem), which gives

$$(5.14) \quad \frac{dF(a)}{da} = -\frac{1}{2}a^3\int_0^{\pi/2}\left[\rho'''\left(\frac{a}{\sin\theta}\right)\frac{1}{a} - \rho''\left(\frac{a}{\sin\theta}\right)\frac{\sin\theta}{a^2}\right] da =$$

$$= -\frac{1}{2}a^2\int_0^{\pi/2}\rho'''\left(\frac{a}{\sin\theta}\right)d\theta + \left[-\frac{1}{2}a\rho''\left(\frac{a}{\sin\theta}\right)\cos\theta\right]_0^{\frac{\pi}{2}}$$

$$-\frac{1}{2}\int a^2\rho'''\left(\frac{a}{\sin\theta}\right)\frac{\cos^2\theta}{\sin^2\theta}\,d\theta =$$

$$= -\frac{1}{2}a^2\int_0^{\pi/2}\rho'''\left(\frac{a}{\sin\theta}\right)\frac{\sin^2\theta + \cos^2\theta}{\sin^2\theta}\,d\theta =$$

$$= -\frac{1}{2}a^2\int_0^{\pi/2}\rho'''\left(\frac{a}{\sin\theta}\right)\frac{d\theta}{\sin^2\theta} \quad .$$

The relation $\rho''(\infty) = 0$ which we have used, follows from the subsequent lemma 5c. Our theorem is thus completely proved.

Lemma 5b.

$$\int_0^\infty u^2 r''(u)\, du = \frac{8}{3\pi} . \tag{5.15}$$

Proof. Integrating by part, we have

$$\int_0^\infty u^2 r''(u)\, du = - \int_0^\infty 2\, u\, r'(u) du =$$

$$= \frac{8}{\pi} \int_0^1 u \sqrt{1 - u^2} = \frac{8}{3\pi} .$$

Lemma 5c. Any correlation function $\rho(u)$, which fulfills the conditions (d), (e) and (f) of Theorem 5.1, has the following properties:

$$\lim_{a \to 0} a^3 \int_0^\infty \frac{1}{u^2} r''\left(\frac{a}{u}\right)\rho''(u)\, du = \tag{5.16}$$

$$= \lim_{a \to \infty} a^3 \int_0^\infty \frac{1}{u^2} r''\left(\frac{a}{u}\right)\rho''(u) du = 0.$$

$$\lim_{a \to 0} a\, \rho'(a) = \lim_{a \to \infty} a\, \rho'(a) = 0. \tag{5.17}$$

If, moreover, $(\rho''(u))/u$ is nondecreasing, then

$$\lim_{a \to 0} a^2\, \rho''(a) = \lim_{a \to \infty} a^2\, \rho''(a) = 0. \tag{5.18}$$

Proof. As $\rho(u)$ is convex, $-\rho'$ is nonnegative and nonincreasing, and we have

$$0 \leqq -\frac{1}{2} a \rho'(a) \leqq -\int_{a/2}^{a} \rho'(u)du = \rho(\tfrac{a}{2}) - \rho(a).$$

As $\rho(u)$ is continuous in points 0 and ∞, (5.17) is clear.

Now, if (5.8) holds true, the function $(\rho''(u))/u$ is nonnegative and nonincreasing, so that

$$(5.19) \qquad 0 \leqq \frac{1}{3} a^2 \rho''(a) = \frac{1}{3} a^3 \frac{\rho''(a)}{a} \leqq a^2 \int_{a/2}^{a} \frac{\rho''(u)}{u} du \leqq$$

$$\leqq 2a \int_{a/2}^{a} \rho''(u)du = 2a \left[\rho'(a) - \rho'(\tfrac{1}{2} a) \right]$$

where the expression converges to 0 if $a \longrightarrow 0$, or $a \longrightarrow \infty$.

The same consideration will be used in proving (5.16). In view of the condition (f), we have

$$(5.20) \quad 0 \leqq \frac{1}{3} a^3 \int_0^{\infty} \frac{1}{u^2} r''(\tfrac{a}{u}) \rho''(u)du \leqq$$

$$\leqq \int_{a/2}^{a} \int_0^{\infty} \frac{\mu^2}{u^2} r''(\tfrac{\mu}{u}) \rho''(u)du\, d\mu =$$

$$= \int_0^{\infty} \left[\int_{a/2}^{a} \frac{\mu^2}{u^2} r''(\tfrac{\mu}{u}) d\mu \right] \rho''(u)du =$$

$$= \int_0^{\infty} \left[\int_{a/2u}^{a/u} u^2 r''(u)du \right] u\rho''(u)du$$

where, in view of (5.15)

$$\int_{a/2u}^{a/u} u^2 r''(u)du \begin{cases} \leqq \dfrac{8}{3\pi} & \text{if } \dfrac{a}{u} > 0 \\ < \dfrac{8}{3\pi} \varepsilon & \text{if } \dfrac{a}{u} \leqq \varepsilon \\ = 0 & \text{if } \dfrac{a}{u} > 1. \end{cases}$$

Consequently, on the one hand,

$$(5.21) \qquad \int_0^\infty \left[\int_{a/2u}^{a/u} u^2 r''(u)du\right] u\rho''(u)du \leqq \frac{8}{3\pi}\left[\varepsilon + \int_0^{a/\varepsilon} u\rho''(u)du\right]$$

and, on the other hand,

$$(5.22) \qquad \int_0^\infty \left[\int_{a/2u}^{a/u} u^2 r''(u)du\right] u\rho''(u)du \leqq \frac{8}{3\pi}\int_a^\infty u\rho''(u)du.$$

The inequalities (5.20) together with the inequality (5.21) or (5.22) prove the relation (5.16) for $a \longrightarrow 0$ or $a \longrightarrow \infty$, respectively. (Notice that $u\rho''(u)$ is integrable with $\int_0^\infty u\rho''(u)du = 1$.)

<u>Example</u> 5.1. The convex correlation function e^{-cu} has a negative third derivative and therefore fulfills the condition (5.8). Hence, it admits the representation (5.1), where the spectral density is given (5.9).

<u>Example</u> 5.2. The convex correlation function

$$\rho(u) = \frac{2}{\sqrt{2\pi}} \int_u^\infty e^{-\frac{u^2}{2}} du$$

also admits the representation (5.1), since $(\rho''(u))/u = (1/\sqrt{2\pi})\, e^{-(u^2/2)}$ is a nonincreasing function of u.

<u>Example</u> 5.3. The linear convex correlation function

$$\rho(u) = 1 - u \qquad \text{if } u \leqq 1$$
$$= 0 \qquad \text{if } u \geqq 1$$

has a discontinuous first derivative, and therefore, does not admit the representation (5.1). It may be shown, however, that $\rho(u)$ is not a planner isotropic correlation function. Actually, considering a square net of points with coordinates

$$p_{ij} = \left(\frac{i}{\sqrt{2}}, \frac{j}{\sqrt{2}}\right) \quad 0 \leqq i, j \leqq n - 1$$

we have

$$\sum_{i,j=0}^{n} \sum_{i',j'=0}^{n} \rho(|p_{ij} - p_{i'j'}|)(-1)^{i+j+i'+j'} =$$

$$= n^2 - 4\left(1 - \frac{1}{\sqrt{2}}\right) n(n - 1)$$

which is negative for a sufficiently large n.

Example 5.3 shows that the class of planner isotropic convex correlation functions is smaller than the class of linear convex correlation functions. One might suspect that all plannar isotropic correlation functions are expressible in the form (5.1). This is, however, disproved by the following:

Theorem 5.2. There exist isotropic stationary stochastic processes in a plane with correlation function $g(x,y) = f(\sqrt{x^2 + y^2})$ such that $f(u)$, $0 \leqq u < \infty$, is a convex function which cannot be represented in the form

$$f(u) = \int_0^\infty r\left(\frac{u}{a}\right) dF(a), \tag{5.23}$$

where

$$r(u) = \frac{2}{\pi} \{\text{arc cos } u - u\sqrt{1 - u^2}\} \quad \text{if } 0 \leqq u \leqq 1, \tag{5.24}$$

$$= 0 \qquad \text{if } 1 \leqq u < \infty,$$

and $F(a)$ is a distribution function with $F(0+) = 0$.

This theorem follows from the following two lemmas:

Lemma 5d. *The function* $g(x,y) = f(\sqrt{x^2 + y^2})$, *where*

$$f(u) = \frac{2}{\pi}\left(\frac{\pi}{2} - u\right) \qquad \text{for } 0 \leq u \leq 1, \tag{5.25}$$

$$= \frac{2}{\pi}\left\{\left(\arc\sin\frac{1}{u}\right) - \left(u - \sqrt{u^2 - 1}\right)\right\}$$

$$\text{for } 1 \leq u \leq \infty,$$

is a correlation function of a stationary isotropic stochastic process (see Fig. 5.1),

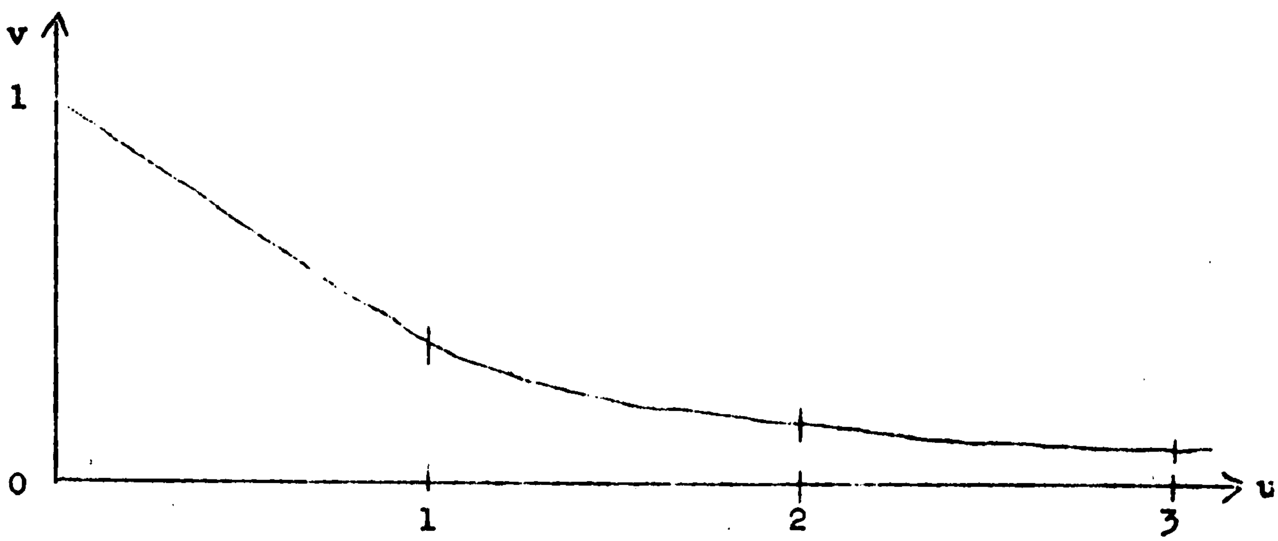

Fig. 5.1. $v = f(u)$, $f(u)$ given by (5.25).

Lemma 5e. *If the function* $f(u)$ *is representable in the form* (5.23), *then* $f''(u) > 0$ *for all* $u > 0$ *with* $u < a$, *where* $F(a) < 1$.

To prove lemma 5d, we consider a linear stationary stochastic process $\eta(t)$ with correlation function $h(t)$ given by

$$h(t) = 1 - |t| \qquad \text{if } |t| \leq 1, \tag{5.26}$$

$$= 0 \qquad \text{otherwise.}$$

Define then a plane stochastic process $\xi(x,y)$ by putting

$$(5.27) \qquad \xi(x,y) = \eta(x \cos \alpha,\ y \sin \alpha),$$

where α is a random variable independent of the process $\eta(t)$ with

$$(5.28) \qquad \Pr\{\alpha < a\} = \begin{cases} 0 & \text{if } a \leqq 0, \\ a/2\pi & \text{if } 0 \leqq a \leqq 2\pi, \\ 1 & \text{if } 2\pi \leqq a. \end{cases}$$

In other words, we first define a plane stochastic process which depends only upon one coordinate and has with respect to it correlation function $h(t)$, and then we randomize the direction. It is seen that $\xi(x,y)$ is an isotropic stationary stochastic process with correlation function $g(x,y) = f(\sqrt{x^2 + y^2})$, where

$$(5.29) \qquad f(u) = \frac{1}{2\pi}\int_0^{2\pi}(1 - h(u \sin \psi))\, d\psi$$

which leads to (5.25).

To prove lemma 5e, we note that the second derivative of a function $f(u)$ given by (5.23) is given by

$$(5.30) \qquad f''(u) = \int_0^{\infty} \frac{1}{a^2}\, r''\left(\frac{u}{a}\right) dF(a),$$

where

$$r''(u) = \frac{4}{\pi}\frac{u}{\sqrt{1 - u^2}} > 0, \qquad \text{if } 0 < u < 1,$$

$$= 0 \qquad \text{if } 1 < u < \infty.$$

This proves lemma 5e.

Now the function $f(u)$ given by (5.25) has a second derivative which vanishes for $0 < u < 1$. This contradicts lemma 5e and thus proves our theorem 5.2.

REFERENCES

[1] W.G. COCHRAN, "Relative accuracy of systematic and stratified random samples for a certain class of populations," The Annals of Mathematical Statistics, (1946), pp. 164-177.

[2] S. DRÁPAL, V. HORÁLEK, Z. REŽNÝ, "Mřížková kvantivní metalografická analysa"(Quantitative metallographic lattice analysis), Hutnické listy, (1957), pp.485-491.

[3] P. FAURE, "Sur quelques résultats relatifs aux fonctions aléatoires stationnaires isotropes introduites dans l'étude expérimentale de certains phénomènes de fluctuations," Comptes Rendus hebdomadaires des seances de l'Académie des Sciences, (1957), pp. 842-844.

[4] Tóth L. FEJES, Lagerungen in der Ebene auf der Ebene auf der Kugel und in Raum., Springer-Verlag, Berlin-Göttingen-Heidelberg, 1953.

[5] D.J. FINNEY, "Field sampling for the estimation of wireworm populations," Biometrics, (1946), pp. 1-7.

[6] J. HÁJEK, "Příspěvky k teorii statistického odhadu, kandidátská disertace"(Some contributions to the theory of estimation), Thesis, Praha, 1955.

[7] J. HÁJEK, "Optimum strategy and other problems in probability sampling," Časopis pro pěstování matematiky, (1959), pp. 387-423.

[8] A. HEPPES, Mehrfache gitterförmige Kreislagerungen in der Ebene," Acta Mathematica Academiae Scientiarum Hungaricae, (1959), pp. 141-148.

[9] A. HUSU, "About certain functionals on stochastic fields," Vestnik Leningradskogo Universiteta, (1957), pp. 37-45.

[10] W.G. MADOW, och L.H. MADOW, "On the theory of systematic sampling," The Annals of Mathematical Statistics, (1944), pp. 1-24.

[11] P.C. MAHALANOBIS, "On large-scale sample surveys," Philosophical Transactions of the Royal Society of London, Series B, Vol. 231, (1944), pp. 329-451.

[12] B. MATÉRN, "Metoder att uppskatta noggrannheten vid linje- och provytetaxering," (Methods of estimating the accuracy of line and sample plot surveys), Meddelanden från Statens skogsforskningsinstitut, ("Reports from the Forest Research Institute of Sweden"), Stockholm, 1947.

[13] B. MATÉRN, "Spatial Variation," (to be published in Meddelanden från Statens skogsforskningsinstitut ("Reports of the Forest Research Institute of Sewden")), Stockholm, 1960.

[14] J.G. OSBORNE, "Sampling errors of systematic and random surveys of cover-type area," Journal of the American Statistical Association, (1942), pp.256-270.

[15] M.H. QUENOUILLE, "Problems in plane sampling," The Annals of Mathematical Statistics, (1949), pp.355-375.

[16] M. SAVELLI, "E'tude expérimentale du spectre de la transparence locale d'un film photographique uniformément impressionné," Comptes Rendus hebdomadaires des séances de l'Academie des Sciences, (1957), pp.871-873.

[17] H.F. SMITH, "An empirical law describing heterogeneity in the yields of agricultural crops," The Journal of Agricultural Science, (1938), pp. 1-23.

[18] SULANKE, Thesis, (1959).

[19] P. WHITTLE, "On stationary processes in the plane," Biometrika, (1954), pp. 434-449.

[20] P. WHITTLE, "On the variation of yield variance with plot size," Biometrika, (1956), pp. 337-343.

[21] S.D. WICKSELL, "The corpuscle problem," First memoir, "Case of Spherical Corpuscles," Biometrika (1925), pp. 84-99; Second memoir, "Case of ellipsoidal corpuscles," Biometrika, (1926), pp. 152-172.

[22] R.M. WILLIAMS, "The variance of the mean of systematic samples," Biometrika,(1956), pp. 137-148.

[23] S. ZUBRZYCKI, "O szacowaniu parametrów złóż geologicznych," (on estimating gangue parameters), Zastosowania Matematyki, (1957), pp. 105-153.

[24] S. ZUBRZYCKI, "Remarks on random, stratified and systematic sampling in a plane," Colloquium Mathematicum, (1958), pp. 251-264.